Person Identification
from Biological Motion

Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
in der
Fakultät für Psychologie
der
RUHR-UNIVERSITÄT BOCHUM

vorgelegt von
Cord Westhoff

April 2005
...wherein, by the colour of his beard,
the shape of his leg,
the manner of his gait,
the expressure of his eye, forehead, and complexion,
he shall find himself feelingly personated.

(William Shakespeare – Twelfth Night)
# Table of Contents

## ABSTRACT

## 1 GENERAL INTRODUCTION

## 2 THEORETICAL BACKGROUND

2.1 THE WORK OF GUNNAR JOHANSSON 4

2.2 METHODS FOR THE CREATION OF POINT-LIGHT STIMULI 6

2.3 THE PERCEPTION OF BIOLOGICAL MOTION 8

2.3.1 Local and global processes in biological motion perception 8

2.3.2 The inversion effect in the perception of biological motion 12

2.3.3 The role of attention and prior experience 15

2.3.4 The perception of biological motion in infancy 16

2.3.5 Summary of the results 17

2.4 THE INFORMATION CONTENT OF BIOLOGICAL MOTION 18

2.4.1 Discrimination of actions 18

2.4.2 Perception of emotions 19

2.4.3 Perception of dynamics 20

2.5 STRUCTURAL AND KINEMATIC INFORMATION 22

2.6 THE ROLE OF THE VIEWPOINT 24

2.7 PERSON IDENTIFICATION FROM BIOLOGICAL MOTION 25

2.7.1 The recognition of gait patterns 25

2.7.2 Identification studies in other domains 30

2.7.3 The understanding of actions 32

2.8 SUMMARY 34

2.9 OBJECTIVES OF THE CURRENT WORK 35

## 3 GENERAL METHODS

3.1 DATA COLLECTION 38

3.2 DECOMPOSITION OF THE DATA 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td><strong>STUDY 1</strong></td>
<td>42</td>
</tr>
<tr>
<td>4.1</td>
<td><strong>INTRODUCTION</strong></td>
<td>42</td>
</tr>
<tr>
<td>4.2</td>
<td><strong>METHODS</strong></td>
<td>42</td>
</tr>
<tr>
<td>4.3</td>
<td><strong>RESULTS</strong></td>
<td>47</td>
</tr>
<tr>
<td>4.4</td>
<td><strong>DISCUSSION</strong></td>
<td>53</td>
</tr>
<tr>
<td>5</td>
<td><strong>STUDY 2</strong></td>
<td>56</td>
</tr>
<tr>
<td>5.1</td>
<td><strong>INTRODUCTION</strong></td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td><strong>EXPERIMENT 1</strong></td>
<td>57</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Methods</td>
<td>57</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Results and Discussion</td>
<td>60</td>
</tr>
<tr>
<td>5.3</td>
<td><strong>EXPERIMENT 2</strong></td>
<td>70</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Methods</td>
<td>70</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Results and Discussion</td>
<td>71</td>
</tr>
<tr>
<td>5.4</td>
<td><strong>SUMMARY AND DISCUSSION OF STUDY 2</strong></td>
<td>77</td>
</tr>
<tr>
<td>6</td>
<td><strong>STUDY 3</strong></td>
<td>80</td>
</tr>
<tr>
<td>6.1</td>
<td><strong>INTRODUCTION</strong></td>
<td>80</td>
</tr>
<tr>
<td>6.2</td>
<td><strong>GENERAL METHODS</strong></td>
<td>81</td>
</tr>
<tr>
<td>6.3</td>
<td><strong>EXPERIMENT 1</strong></td>
<td>84</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Methods</td>
<td>84</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Results and Discussion</td>
<td>86</td>
</tr>
<tr>
<td>6.4</td>
<td><strong>EXPERIMENT 2</strong></td>
<td>89</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Methods</td>
<td>89</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Results and Discussion</td>
<td>90</td>
</tr>
<tr>
<td>6.5</td>
<td><strong>EXPERIMENT 3</strong></td>
<td>94</td>
</tr>
<tr>
<td>6.5.1</td>
<td>Methods</td>
<td>95</td>
</tr>
<tr>
<td>6.5.2</td>
<td>Results and Discussion</td>
<td>96</td>
</tr>
<tr>
<td>6.6</td>
<td><strong>EXPERIMENT 4</strong></td>
<td>98</td>
</tr>
<tr>
<td>6.6.1</td>
<td>Methods</td>
<td>99</td>
</tr>
<tr>
<td>6.6.2</td>
<td>Results and Discussion</td>
<td>99</td>
</tr>
<tr>
<td>6.7</td>
<td><strong>SUMMARY AND DISCUSSION OF STUDY 3</strong></td>
<td>102</td>
</tr>
<tr>
<td>7</td>
<td><strong>GENERAL DISCUSSION</strong></td>
<td>107</td>
</tr>
</tbody>
</table>

**REFERENCES** 118
Abstract

Human observers’ ability to identify persons from point-light displays of their gait patterns was examined in three studies. Point-light representations were created by recording individual locomotion patterns with a motion capture system and by subsequently decomposing the data using Fourier analysis. In the first study, participants learned to name seven previously unknown male point-light walkers. Throughout the experiment, individual parameters were removed from the displays and replaced by averaged values. It was shown that body shape and gait frequency, but not size, play significant roles in the discrimination of individual persons shown as point-light displays. The second study focused on the kinematic details of the gait patterns. Again, observers learned to name seven unknown male walkers. Afterwards, individual harmonics (Experiment 1) and amplitude- and phase-spectra (Experiment 2) of the gait patterns were normalized. Only the first and second harmonics were found to be of relevance for the identification task. Moreover, normalization of amplitude- as well as of phase-spectra significantly impaired the recognition of the individual walkers, although removing information about the amplitudes resulted in a much stronger deterioration. When the displays were rotated, a significant transfer to novel viewpoints was found, although best performance resulted from identical training- and test-viewpoints. The third study examined the effect of the familiarity with gait patterns on a number of perceptual tasks. Two groups of subjects that were known to each other served as models and as participants in the test sessions. The experiments contained the detection of a walker in a scrambled mask (Experiment 1), the discrimination between forward and reversed animations (Experiment 2), and the adjustment of the original gait frequency (Experiment 3). No differences were found between point-light displays which were based on the subjects’ own gait patterns, on those of their group members, or on those of the participants of the other group. Finally, the subjects had to identify the individual group members from their point-light depictions (Experiment 4). A relatively low performance, compared to earlier studies on this topic, was found. Implications for the perception of biological motion, as well as possible applications of the findings, are discussed.
1. General Introduction

Many species are highly sensitive to motion patterns generated by other living organisms. The detection and the identification of different kinds of motion enable animals to evaluate potential threats or possible prey and thus enhance the probability of survival. In a similar way the recognition of animate motion can help humans to initiate contact with their conspecifics.

Probably the most important aspect of motion recognition is its visual perception, compared to tactile or auditory perception, because the movements contain a multitude of information, communicated through the visual domain. It is a well supported fact that the interaction with other persons is not only based on verbal communication, but that it is also strongly supported by a non-verbal exchange. Body language, even very subtle movements, can convey emotions, intentions, hidden meanings and many other hints on factors of social interaction. The visual system is highly adept on detecting and interpreting these signs, but before we can engage in close interaction, a previous step is to identify persons who we want to get into contact with.

We can often recognize a well familiar person from a distance, even before we are able to perceive detailed features such as the face, clothing, hair, or colour of skin. Personal identity can be revealed through very specific motions of the body. Probably the most frequently observed motion pattern is locomotion, hence other persons can often be identified by the way they walk. Whereas it is commonly experienced that we can recognize other persons by their ways of locomotion, it is still unknown how we accomplish this extraordinary task and which cues have a share in person identification. The present work contributes to this question by examining the factors that are used to discriminate individual persons. The specific roles of structural as well as of kinematic cues, which can be separated by advanced techniques, are in the focus of attention. Moreover, the effects of the familiarity of gait patterns on several perceptual tasks are analyzed.

The following chapter will provide an overview of the findings about the perception of biological motion and specify the objectives of this work. The
general method that is used for the creation of the stimuli will be described in Chapter 3. This will include a depiction of the motion capture technique, as well as a description of the decomposition of the data into structural and kinematic components. Chapters 4-6 contain the studies that were conducted to examine different aspects of person identification from biological motion. Whereas the first study examined the role of a walker's size, body structure, and gait frequency on his recognition, the second study focused on the kinematic details of a point-light representation. The third study examined the influence of the familiarity with a gait pattern on a variety of perceptual tasks. The final chapter will provide a general discussion of the results, as well as suggestions about the practical implications of the findings.
2. Theoretical Background

The first part of this chapter is dedicated to the pioneering work of Gunnar Johansson, who initiated an immense amount of research on what he called “biological motion”, namely the characteristic locomotion patterns of living organisms. After a description of Johansson’s methods and findings, other techniques to collect or to synthesize human motion patterns will be presented. Due to technical developments over the last decades, more sophisticated methods for the creation of biological motion stimuli have become available. All studies on biological motion rely on video-based methods, computer-generated stimuli, or on a combination of both. An important aspect in the description of these techniques is the ecological validity of the point-light stimuli under the aspect of distal and proximal approaches, that is, of the original recordings and the displays that appear on the screen.

The next section will describe some of the basic principles of how the visual system accomplishes the task of organizing complex stimuli of biological motion into a coherent structure. These principles comprise local as well as global analyses. Moreover, the role of attention and prior expectations, as well as the development of these processes in infancy are described.

Having established these fundamental capabilities, I will summarize the findings about which kinds of information are contained in these point-light displays. This will include both, information about personal properties, such as a person’s gender or emotions, as well as information about the environment, especially the dynamic principles which act on a person’s movements. The next parts depict the decomposition of point-light displays into structural and kinematic components, and the role of the viewpoint under which an object or event is perceived.

This will be followed by a description of the studies that have examined person identification from biological motion as well as the effect of the familiarity with a motion pattern on the recognition of individual persons. Furthermore, some examples will be given about person identification from other domains, like speech perception or the trajectories of drawn characters.
2.1 The work of Gunnar Johansson

A vast amount of research activity about the perception of animate motion was initiated by the pioneering work of Gunnar Johansson (1973, 1976), who termed biological motion as “motion patterns characteristic of living organisms in locomotion”. Whereas Johansson had studied the perception of synthetic, abstract events before (e.g., Johansson, 1950, 1958, 1964), he turned to investigate natural events. Most important, he brought the distal stimulus into focus, that is, the original, unaltered event. He studied the perception of biological motion patterns after all other information about the actor or the environment had been subtracted from the displays. The first method to accomplish this was to attach flashlight bulbs to the main joints of a walking actor who was filmed from the side. This arrangement was found to have serious drawbacks, especially with the wire connections of the flashlights, leading to a limited degree of freedom for the actors. In the second method therefore “reflex patches” were attached to the joints, which reflected the light of searchlights. After the motions were recorded onto a video tape, a TV monitor was adjusted with high contrast and low brightness, so that only the reflections of the patches could be seen on the screen. No other pictorial information about the actor remained in the display. Although this general technique was not new (see Marey, 1895/1972), Johansson was the first who applied it to study the perception of complex events.

Johansson (1973) described that all observers immediately reported seeing a human figure in motion when the video recordings were shown to them, but that this was never the case with a static picture of the lights. Obviously, observers used the contingent motion of the separated dots to infer an underlying human structure from an otherwise meaningless visual pattern. Moreover, observers had no problems to identify the human structure if the viewpoint deviated from the sagittal view of the person. In these cases, due to
the projection of the lights onto the screen, the constant distances of the lights on the actor’s body continuously changed their projected distances in the display, but the perception of a rigid structure did not change. Even if only five lights representing the hip and the legs of the actor were shown, observers reported seeing a walking human figure. This percept was not affected by subtracting the common translatory component of the overall pattern, such that the figure appeared as if walking on a treadmill, or by adding an additional circular motion to the trajectory of each element. Moreover, Johansson (1976) demonstrated that a presentation time of 200 ms was sufficient to recognize the lights as representing a human figure. A duration of 400 ms was sufficient to discriminate various kinds of motion patterns conveyed by the actors, and to separate real motions from the movements of puppets.

Johansson (1950) had already suggested the principle of “perceptual vector analysis” which accounted for the perception of visual motion. It consisted of a set of simple rules for the mathematical vector analysis of the stimulus. Probably the most important principle is that a basic goal of the visual system is to abstract equal simultaneous motion vectors from a set of proximal elements, which are perceived as a common, unitary motion against its environment, similar to the Gestalt law of common fate. In the context of biological motion we are dealing with a hierarchical structure of common motions. At the top of this hierarchy is the motion of the trunk (i.e., of the shoulders and hips). The motions of these joints are common to all other elements of the walking figure. Moreover, the motions of the elbows and the knees are pendulum motions relative to this basic frame of reference. Furthermore, the movements of the wrists and ankles in turn are pendulum motions around the elbows and knees. Support for this model came from the findings that the subtraction of the translatory motion as well as the addition of other common components does not change the overall percept of a walking figure. Apparently, these common motion vectors are filtered by the visual system, leaving only the trajectories of the individual joints.
2.2 Methods for the creation of point-light stimuli

Since Johansson’s inspiring work more sophisticated techniques have been used to create point-light displays of biological motion. Cutting (1978a, 1978b) developed a program to synthesize human motion as point-light displays. Measurements from samples of individual walkers were taken to estimate structural and kinematic parameters. All joint movements were pendular or ellipsoidal, and the motions were chosen such as they were “reasonably good-looking” or “fitted by eye”. The goal was “to find minimal rules for the synthesis of certain kinds of gaits.” (Cutting, 1978a, p. 397).

Runeson (1994) pointed out that Cutting’s algorithm (1978a, 1978b) resembled only a proximal stimulus, because it was not based on any real recordings of animate motion. Thus, although many studies have adapted a version of this algorithm (e.g., Bertenthal, Proffitt, & Cutting, 1984; Neri, Morrone, & Burr, 1998; Pinto & Shiffrar, 1999; Thornton, Rensink, & Shiffrar, 2002; Thornton & Vuong, 2004), the findings of those studies are only of a questionable validity to natural settings. In particular, the point-light walker created by Cutting lacks the dynamic forces acting on the body. This is especially evident in the very smooth movements of the legs and the feet, as well as in the lack of bounce in the overall body. Dynamic factors such as masses or gravity are not considered in the algorithm created by Cutting (1978a, 1978b).

Proffitt, Bertenthal, and Roberts (1984) examined the role of occlusion on the perception of this type of synthesized walker. They adapted Cutting’s algorithm, such that the point-lights of the elbow, wrist, knee, and ankle on the far side of the body disappeared as they would if they were occluded by the body in a real situation. When observers were asked what they had seen after they were shown the displays for 1.5 min, 83% reported having seen a person walking in the occlusion condition. Surprisingly, only 33% reported to see a human figure when there was no occlusion in the display. The occlusion condition resembles Johansson’s video displays, but even after 90 seconds not all observers perceived a human figure. These results are in striking contrast to the results of Johansson (1976), who found an immediate and exclusive
perception of a human walker from his video recordings after only a fraction of a second. This indicates a lack of naturalness of the algorithm as advanced by Cutting (1978a, 1978b) as well as a high sensitivity of the visual system for deviations from natural motion.

Grossman and Blake (1999; see also Grossman, Blake, & Kim, 2004) applied a method that was based on the distal stimulus. The original recordings of an actor were imported on a frame-by-frame basis into a computer, where the dot positions were coded into successive matrices. This left the original stimulus intact but allowed its subsequent modification for comparative purposes. The main purpose for the application of this technique was to create a mask of noise dots which is based on the original trajectories of the markers on the actor’s body.

Another method that relies on video processing software is described by Thomas and Jordan (2001). The authors pointed out the importance of comparing point-light animations and fully illuminated displays which are both based on the same piece of video footage. They argue that only in this case it is possible to precisely match these two stimuli, which would otherwise be impeded by a natural variance between different occasions of recording. Thomas and Jordan (2001) therefore describe a technique that is based on the video of an actor wearing reflective markers. This single video can subsequently be edited by the use of video tools to retain either only a display of the markers or a video of the actor without the reflective markers. Although this technique was originally developed for facial motions, it should be equally applicable to the production of whole-body movements.

The methods used by Grossman and Blake (1999) and by Thomas and Jordan (2001) both rely on the original stimulus and thus bear a high ecological validity. On the other hand, a serious limitation is the restriction to a single viewpoint, as both techniques use video cameras to create the basic video footage. The technique that was used for the creation of the stimuli in the present work also uses the original recordings of the actors, but it is much more flexible regarding the viewpoint of the presentation. A set of reflective markers is attached to the body of the respective actor, and the trajectories of these markers are digitally recorded by nine cameras, positioned around the actor,
with a high spatial and temporal accuracy. From the data of these cameras, the 3-dimensional position of every marker can be reconstructed for every point in time. This procedure allows the preservation of the original, individual movements, and the presentation of the proximal stimulus on the screen is not restricted to a single viewpoint, but it can be rotated around all three axes. A more detailed description of this method and the subsequent decomposition of the data are given in Chapter 3.

2.3 The perception of biological motion

2.3.1 Local and global processes in biological motion perception

A large body of research indicates that the perception of biological motion depends on the analysis of the local motion of the individual dot-trajectories, as well as on the analysis of the global figure. Although there are no clear definitions as to how these two types of information processing are to be separated, it is generally agreed that local analyses of biological motion depend on the computation of individual points (representing the joints of the body) or point pairs (rigid connections between these joints), whereas global mechanisms process information across larger spatiotemporal intervals. Local motion analyses proceed only over a short time interval of up to 50 ms (Baker & Braddick, 1985). A separation of the individual frames of an animation by more than 50 ms should therefore lead to a disruption of these processes.

Support for the importance of local processes comes from studies of Cutting, Moore, and Morrison (1988) and from Mather, Radford, and West (1992). The former authors showed that the recognition of the walking direction of a point-light stimulus is significantly impaired when it is camouflaged by a mask of dots that contain motion trajectories identical to those of the dots of the walker (scrambled walker mask), thus rendering the information of individual point-lights uninformative.

Mather et al. (1992) found that the correct identification of the walking direction depends on the interval between individual frames (Inter-Stimulus-
Interval, ISI) of a display. When an ISI of 60 ms or longer was inserted between two consecutive frames, the performance in the task dropped to chance level. This is in accordance with the assumption that low-level processes can only integrate information over a time span of 50 ms. Taken together, these two results hint to a local analysis, because spatial and temporal disruption of these processes impairs the perception of walking direction.

Mather et al. (1992) also found that the omission of the extremities of the point-light walker (wrists and ankles) significantly impaired direction discrimination, whereas this was not the case for the omission of middle (elbows, knees) or central (shoulders, hips) elements. A study by Pinto and Shiffrar (1999), on the other hand, showed that the extremities are not necessary for the detection of biological motion per se. They replicated the experiment of Mather et al. (1992) with a detection task. Under such conditions the omission of middle and central elements, but not the omission of the extremities, impaired detection performance, suggesting that the relative influence of the body parts depends on the task at hand. Whereas the discrimination of walking direction is apparently enhanced by the presence of the body’s extremities, the detection of biological motion seems to depend on the central elements, which is in accordance to computational models as suggested by Hoffman and Flinchbaugh (1982) or Webb and Aggarwal (1982). These authors suggested algorithms that test for pairs of points with local rigidity. The detection of a figure is enhanced by the presence of segments which are connected via shared joints. If these shared joints, for example, the hip, are removed from the displays, the detection of connected segments is impaired.

Moreover, Pinto and Shiffrar (1999) suggest that the visual system searches for characteristic elements or subconfigurations of a target to detect biological motion. This was concluded from the finding that a random configuration of the limbs impaired the detection of biological motion, whereas this was not the case when coherent subconfigurations of a walker (contralateral limbs, ipsilateral limbs, or diagonal limbs) were shown. Specific organization principles, like dynamic symmetry of the limbs or the principal axis of the body, did not unambiguously lead to differential enhancement of the
detection of a walker in visual noise, or to the general perception of a human structure. These results were taken as evidence that the recognition of biological motion is not solely based on local analyses, but that the global perception of a figure (or its parts) does also play an important role.

Thornton, Pinto, and Shiffrar (1998) found further evidence that local as well as global processes contribute to the recognition of a human form. Point-light displays were shown in a random mask (i.e., the positions of the noise dots did not follow a coherent pattern but randomly varied on a frame-by-frame basis) either for a duration of 20 frames, or for a duration of 80 frames, each one separated by an ISI ranging from 0 to 120 ms. Whereas performance remained relatively stable over variable ISIs in the 80-frame condition, detection of the walker diminished significantly in the 20-frame condition at an ISI of 60 ms or longer. A purely local or global mechanism should have produced the same results in both conditions, independent of the overall duration of the display. This indicates that the visual system is able to integrate information under longer ISIs, given a sufficient amount of processing time. In a second experiment, the same design was repeated with a scrambled mask to eliminate the use of low-level spatial processes. Performance in the 20-frame condition was above chance only for very short ISIs. In the 80-frame condition, on the other hand, there was a gradual drop of performance, reaching chance level only at long ISIs above 100 ms. These findings indicate that low-level spatial analyses are performed as long as local motion cues are available to the visual system. On the other hand, the fact that performance above chance level was found for ISIs of more than 50 ms suggests that point-light stimuli can be tracked by a high-level system over extended spatio-temporal intervals.

The global perceptual process even tolerates spatial perturbations when the dots are not placed at the main joints of the walker but at in-between positions (Cutting, 1981; Dittrich, 1993). Beintema and Lappe (2002) extended these findings by showing that the recognition of a walking person is also possible when the positions of the dots randomly vary on a frame-by-frame basis. In contrast, temporal perturbations, created by phase-shifting the relations of the dots, significantly impair the recognition of the structure in a simultaneous mask (Bertenthal & Pinto, 1994). By phase-shifting the point-light
representation of a walker the spatial positions of the dots are left unchanged but they vary in their starting positions in the walk cycle. For example, the motions of the left leg and the right arm, moving in approximate synchrony in a normal gait, are randomly shifted against each other. Phase-shifts can be detected by observers across a variety of conditions, for example, contrast polarity and spatial frequency of the point-lights, or first- vs. second-order motion (Ahlström, Blake, & Ahlström, 1997).

Further evidence for a global perception of biological motion, but using a paradigm of apparent motion, comes from studies of Shiffrar and colleagues (Shiffrar & Freyd, 1990, 1993; Shiffrar, Lichtey, & Heptulla Chatterjee, 1997). Under conditions of apparent motion, two stationary targets shown on consecutive frames are perceived as one moving target. Observers typically report the target as moving along the shortest possible path between the two positions. In contrast, Shiffrar and Freyd (1990, 1993) found evidence for violations of this rule with photographs of human bodies. When observers saw photographs of a human with a differently positioned limb on the two pictures, they only reported seeing the limb moving along the shortest path (even if this motion violated physical constraints) under conditions of very short ISIs. With sufficiently long ISIs however, they reported the perception of the limb moving along an extended path but according to physical constraints of the body. Moreover, the necessary processing time for the perception of appropriate motions was related to the length of the motion path required to satisfy biomechanical limitations. Non-biological objects (clocks, rectangles) were always seen as moving along the shortest possible paths. The perception of biological motion, therefore, is apparently based on a high-level representation of physical constraints. These concepts are activated if the visual system is given a sufficient amount of processing time.

Taken together, the studies mentioned so far indicate that the perception of biological motion relies on local as well as on global processes. Local analyses, relying on the trajectories of individual points or point-pairs, contribute to the recognition of a human structure. If these low-level cues are disrupted by using a scrambled mask or by the insertion of an interval of more than 50 ms between individual frames, the perceptual process can be severely impaired.
However, a supposed high-level process can integrate information over longer intervals, given a sufficient amount of time.

2.3.2 The inversion effect in the perception of biological motion

Further contributions to the discussion about local and global analyses come from the inversion effect of point-light displays. Turning the display upside down severely diminishes the recognition of a human walker as such, as well as of the detection of form-from-motion under conditions of visual noise (e.g., Bertenthal & Pinto, 1994; Dittrich, 1993; Pavlova, 1989; Pavlova, Lutzenberger, Sokolov, & Birbaumer, 2004; Pinto & Shiffrar, 1999; Sumi, 1984). In an inverted display, the hierarchical relations between the dots are preserved, but this is obviously not sufficient to ensure detection of the target structure.

Using Johansson’s (1973, 1976) original movie clips, Sumi (1984) first reported that an upside-down presentation of these displays disrupted the ability of most observers to correctly recognize an inverted human figure. Instead, about half of the subjects reported having seen an upright human figure performing movements like jogging, punching and ducking, walking with tottering steps etc. However, these responses might have been facilitated by the restriction of the answers to biological/nonbiological or human/nonhuman categories. Furthermore, Sumi (1984) did not subtract the translational component from the displays.

Even prior information about the orientation of the stimulus does not lead to a reliable recognition of a point-light walker or of an animated dog (Pavlova & Sokolov, 2003). On the other hand, the inverted stimuli become better recognizable when a moving or a static background line is inserted into the animation. These lines might serve as an additional frame of reference for the veridical orientation of the displays.

Display inversion also impairs the recognition of the gender of a point-light walker (Barclay, Cutting, & Kozlowski, 1978) or of the type of an action. Dittrich (1993) found that some simple or highly familiar actions can be
discriminated better than others when shown upside down. He interpreted this as evidence for a knowledge-based, top-down process.

The impaired recognition of inverted displays is further underlined by brain imaging data from Grossman and Blake (2001). These authors found that viewing an upright point-light walker activates the posterior superior temporal sulcus (STS) about twice as much as when the same display is turned upside down. The STS was found to be of major importance in the processing of biological motion in a number of studies (e.g., Beauchamp, Lee, Haxby, & Martin, 2003; Grossman et al., 2004; Pelphrey, Mitchell, McKeown, Goldstein, Allison, & McCarthy, 2003; Servos, Osu, Santi, & Kawato, 2002; Vaina, Lemai, Bienfang, Choi, & Nakayama, 1990; Wheaton, Thompson, Syngeniotis, Abbott, & Puce, 2004). The STS is not simply activated by viewing complex motion patterns, but it responds preferentially to articulated biological motion (Pelphrey et al., 2003). Moreover, an increase in the ability to recognize coherent biological motion in a scrambled mask is accompanied by a significant increase in the activation of the STS (Grossman et al., 2004). These results provide compelling evidence for a connection between the perception of biological motion and an enhanced activation of the STS. Therefore, a diminished activation of this area can be interpreted as an impaired recognition of the stimuli as a type of biological motion.

The inversion effect can be accounted for by an event-from-form model and by an event-from-dynamics model. The former approach suggests that the recognition of an inverted display is impaired because the positions of the joints deviate from their familiar locations as perceived in the everyday environment. Thus, the natural configuration of the stimulus-elements is broken. The event-from-dynamics model, on the other hand, assumes that impaired recognition is caused by the unfamiliar dynamics specified by the kinematics, because the implied gravity is directed upward. Usually, information about familiar configuration and natural dynamics are confounded in studies of biological motion. Shipley (2003) separated these two types of information by recording actors who walked on their hands, thereby inverting the orientation of the figure but leaving the natural dynamics intact. Observers showed an enhanced recognition of point-light figures walking on their hands, as compared to the
upside down presentation of these displays, which inverted gravity but showed the figures in a familiar orientation. It was therefore concluded that the orientation of gravity, but not form, is important for the detection of these point-light stimuli in a scrambled mask.

Further support for this assumption comes from Troje (2004; see also Troje & Westhoff, in preparation). Using a direction discrimination task, the authors found that observers were able to correctly judge the apparent walking direction of scrambled human and animal point-light displays shown from a sagittal view. In contrast, an inversion of the displays strongly impaired the recognition of the walking direction. It was concluded that the inversion effect is not caused by an unusual configuration of the display, because both conditions, inverting and scrambling the stimuli, disturbed the configuration of the displays. Rather, it was suggested that the effect is based on an invalid direction of gravity. The importance of gravity was further supported in a recent fMRI study by Indovina, Maffei, Bosco, Zago, Macaluso, and Lacquaniti (2005). The authors showed that brain regions overlapping with those activated by direct vestibular stimuli were also selectively activated when observers watched animations that were coherent with natural gravity (e.g., insular cortex, temporoparietal junction, or ventral premotor area). It was therefore concluded that mechanisms of physical laws of motions are represented in the human brain.

Troje (2003) found evidence that the perception of biological motion is not impaired by the rotation of the stimuli per se, but by a mismatch between the orientation of the stimulus and the observer. In addition to rotating the stimuli by 90°, he also presented point-light displays to observers who were lying on their sides, and thus had rotated their heads by the same degree. The detection of biological motion was not generally impaired when it was shown under a rotated condition, but only when the orientation of the stimulus did not match the orientation of the observer. This was taken as evidence that the perception of biological motion operates in egocentric coordinates, and that it does not depend on the absolute orientation of the stimulus in an environmental frame of reference.
In summary, a rotation or an inversion of point-light stimuli can severely impair its recognition. Some observers are not even able to recognize a walking human figure if it is presented upside down. This inversion effect is apparently based on the inversion of the underlying gravity of the event, and not necessarily caused by a disturbance of the configural processing.

2.3.3 The role of attention and prior expectations

Several studies focused on the role of attention for the perception of biological motion by the use of distracting tasks or contradicting stimuli. Thornton et al. (2002) argued that if the perception of biological motion relies on high-level demands, the introduction of a secondary task should interfere with recognition and thus impair performance on perceptual tasks. The authors found that direction discrimination of a camouflaged point-light walker was actually impaired when the observers simultaneously performed an additional task, not related to biological motion. Without ISIs to separate the individual frames of the walker animation, this was only the case when the walker was hidden in a scrambled mask. When a random dot mask was used, performance was only impaired by the introduction of ISIs of 40 ms or longer. Thornton and Vuong (2004) found evidence for incidental processing of the direction of a point-light walker by using a flanker-interference paradigm. In a flanker task, observers have to make quick responses to a target item surrounded by additional distractors. The authors showed that response times in a direction discrimination task significantly increased when the target was surrounded by other point-light figures which were walking in the opposite direction, but not by static flankers. This result suggests that the perception of biological motion is partly an automated task, as could be assumed by the expertise of human observers for the recognition of natural locomotion patterns. Nevertheless, a certain degree of directed attention is needed for this task, as shown by Thornton et al. (2002).

Compelling evidence for the influence of prior expectations on depth perception, and thus for the effect of high-level processes, comes from Bülthoff,
Bülthoff, and Sinha (1998). They presented stereoscopic versions of point-light walkers with scrambled depth-order information. As long as the two-dimensional projections of the proximal stimuli were unchanged, recognition of the familiar three-dimensional structure was unaltered, despite the stereoscopic deviations. Moreover, the observers seemed unaware of the anomalies that were introduced by depth-scrambling. Obviously, the low-level process of stereoscopic depth perception can be overridden by previous expectations about the figure to be perceived. Cavanagh, Labianca, and Thornton (2001) suggest that “sprites”, special sets of routines that are attuned to specific characteristics of objects and their changing configurations, are responsible for the relatively effortless perception of a walking human form. Degraded or contradicting information, as used in the study of Bülthoff et al. (1998) can thus be filled with or overridden by representations from memory to support a rich and conclusive percept.

2.3.4 The perception of biological motion in infancy

The recognition of biological motion can be verified at a very early period in life. Bertenthal and colleagues (Bertenthal et al., 1984; Bertenthal, Proffitt, & Kramer, 1987; Booth, Pinto, & Bertenthal, 2002) as well as Fox and McDaniel (1982) assessed the development of infants’ sensitivity to biological motion. It was demonstrated that even by three months of age, children are able to group moving point-light stimuli into a coherent structure. When infants were habituated to a translating canonical point-light walker, a significant amount of response recovery was found when the walker was turned upside down (Bertenthal et al., 1984). No such inversion effect was found when scrambled walkers were presented to the infants. Sensitivity for the local rigidity of point-light walkers was only found in 5-month-old infants, but not in 3-month-olds (Bertenthal et al., 1987). Booth et al. (2002) found evidence that the perception of a global structure develops between the ages of three and five months. They presented point-light depictions of a walker and of a runner to the infants, both containing a symmetrical patterning of the limbs but differing in many other
aspects, like speed or angular velocity and displacement. The 3-month-old infants exhibited a significant response recovery between these two types of stimuli, which was explained by their observation of local differences between the motion patterns. The 5-month-old infants, on the other hand, did not show response recovery, indicating that they paid more attention to the global, symmetrical pattern of the stimuli.

In a recent study using event-related potentials, Hirai and Hiraki (2005) found that 8-month-old infants showed a similar activation pattern like adults in the right hemisphere when they watched biological motion, compared to scrambled motion. Whereas the response to biological motion was significantly greater than the response to scrambled motion in both hemispheres of adults, this was only the case for the right hemisphere of the infants. However, the authors were unable to determine particular areas of activation. They suggest that the neural circuits involved in the perception of biological motion have started to develop by 8 months of age, but that a more precise clarification of the age at which these substrates begin to mature is still needed.

2.3.5 Summary of the results

Taken together, the above mentioned evidence suggests that global information about a walker, as well as analyses of the local trajectories of the point-lights are used to support the detection of a human figure in motion. Neither of the processes can by itself explain the perception of human locomotion, and the relative importance of global and local analyses apparently depends on the specific task at hand. An inversion of the displays severely impairs the recognition of a human figure, as well as recognition of its gender or the action that is performed. This inversion effect is obviously dominated by the inverted direction of gravity, and not by the perception of an unfamiliar human structure. The sensitivity for different parameters of biological motion develops in early infancy; the period between the age of 3 and 5 months seems to be an important phase for the switch of concepts underlying perception.
2.4 The information content of biological motion

Many studies have shown that the visual kinematics of point-light displays convey information about the performed action (Dittrich, 1993; Johansson, 1976; Pollick, Fidopiastis, & Braden, 2001; Sparrow & Sherman, 2001), the gender of the actor (Barclay et al., 1978, Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Sparrow, Shinkfield, Day, & Zerman, 1999; Troje, 2002a, 2002b), traits such as vulnerability (Gunns, Johnston, & Hudson, 2002), emotions (Atkinson, Dittrich, Gemmell, & Young, 2004; Dittrich, Troschianko, Lea, & Morgan, 1996; Heberlein, Adolphs, Tranel, & Damasio, 2004; Pollick, Lestou, Ryu, & Cho, 2002; Pollick, Paterson, Bruderlin, & Sanford, 2001), or the meaning of gestures used in American Sign Language (Poizner, Bellugi, & Lutes-Driscoll, 1981).

The perception of these parameters depends in part on the general ability of information processing. Sparrow et al. (1999) found that adults with mental retardation are impaired in the recognition of gender from point-light displays depicting human activities, but unimpaired in the categorization of the actions themselves. Further evidence for the influence of mental state on the perception of biological motion comes from Blake, Turner, Smoski, Pozdol, and Stone (2003), who showed that the discrimination of canonical and phase-scrambled walkers, but not the recognition of global form, is significantly impaired in children with autism. In contrast, Jordan, Reiss, Hoffman, and Landau (2002) found evidence for a selective sparing of specialized systems for the recognition of biological motion in children with Williams syndrome, in spite of spatial cognitive deficits.

2.4.1 Discrimination of actions

Dittrich (1993) examined the recognition of locomotory, social, and instrumental actions. He found lower accuracy and longer response times for instrumental actions (e.g., ball bouncing, stirring) than for the other two conditions. This is in accordance with results from Shipley and Cohen (2000). These authors
reported that the addition of information about the instrument enhances the perception of the action itself. In this case, observers had to describe what they saw in point-light displays of two basketball players. An additional single point-light representing the ball led to a significant increase in the perception of an animate and coordinated group action. The authors suggest that the additional point-light specified the action in relation to the environmental context.

2.4.2 Perception of emotions

Pollick et al. (2002) found that human observers showed a high sensitivity in discriminating neutral or angry affects from point-light movements (knocking, lifting, waving), in comparison with an artificial neural network. On the other hand, they performed poorly in discriminating the gender of the actors from the same displays. Pollick et al. (2001) suggested a two-dimensional psychological space underlying the recognition of emotions, corresponding to a circumplex structure, with the first dimension appearing as activation and the second as pleasantness. The authors found that observers were able to categorize ten different emotions from point-light displays of an arm performing knocking and drinking movements. The activation dimension, but not pleasantness, could also be derived from scrambled versions of the displays. Therefore, the authors argue that activation is a formless cue which can be perceived from kinematics alone, independent of a coherent shape, whereas the recognition of pleasantness relies on the phase relations between the segments of the limb. Activation was correlated to kinematic properties such as velocity, acceleration, or jerk of the movements.

Dittrich (1996) showed that emotions can also be conveyed through point-light displays of dance movements, although a significantly better performance was found for the recognition of full-body stimuli. This was confirmed by Atkinson et al. (2004), who let actors portray emotions at different levels of exaggeration, with their faces covered. Emotions from full-body movements were more readily identifiable, but even point-light displays could convey sufficient information for the identification of specific emotions.
Exaggeration led to an enhanced overall recognition, as well as to higher ratings of emotion-intensity. Pollick, Hill, Calder, and Paterson (2003) investigated the effects of spatially and temporally exaggerating point-light displays of facial emotions. Whereas manipulating the spatial properties of the displays resulted in enhanced ratings of the intensity of an emotion, changing the duration of an expression only had a small effect on the intensity-ratings.

Bonda, Petrides, Ostry, and Evans (1996) observed an enhanced PET activity in the right superior temporal sulcus and the adjacent temporal structures as well as in the amygdala when observers viewed point-light displays of expressive body movements. Recently, Heberlein et al. (2004) found a dissociation between the judgement of emotions and personality traits in brain-damaged patients. Impairments in judging emotions from point-light displays were connected to lesions in the right somatosensory cortices. On the other hand, a diminished ability to attribute personality traits was found for patients with damage to the left frontal opercular areas.

2.4.3 Perception of dynamics

Observers are not only able to derive properties of the agent of an event depicted as a point-light display, but also about the physical environment and the forces that have an effect on the action. Kinematic information (with the fundamental parameters position and time) is the only cue that is given to an observer when an event is shown as a point-light display. Nevertheless, Runeson and Frykholm (1981, 1983) demonstrated that observers are able to derive the underlying distal dynamics (masses, forces, momentum, etc.) of an action from these proximal kinematic cues. For example, the observers could estimate the weight of a lifted box, or how far a sandbag was thrown. Furthermore, observers could discriminate between a point-light depiction of a real event and a person who was pretending to perform an action. In a recent fMRI study, Grèzes, Frith, and Passingham (2004) found an increased activation in the amygdala, the rostral anterior cingulate cortex, the superior temporal sulcus and the lateral orbitofrontal cortex when observers judged lifting
actions as reflecting a deceptive intention of the actor about the real weight of a box.

Bingham (1987) extended Runeson and Frykholm’s (1981, 1983) findings by showing that point-light displays of one-arm curls are sufficient to enable observers to give accurate estimations of a lifted weight. Thus, even the kinematic pattern of a single-joint movement contains useful information about kinetic causes.

In the studies by Runeson and Frykholm (1981, 1983) observers were presented with standard trials, in which the weight of a lifted box was specified. Therefore, the possibility remained that the observers guessed the individual strength of each lifter in relation to this standard weight and approximated only the relative efforts in the other trials. However, Bingham (1987) found that the scaling in form of a standard weight is not a necessary prerequisite for the estimation of weight. Observers were able to judge the absolute weight of a lifted box over a variety of lifters and weights.

Bingham (1995) proposed that the trajectory form, that is, the path of motion and the variation in the speed profile along the path, is the central piece of information used to identify events. In a recent study, Wickelgren and Bingham (2001) found that 8-month-old infants are sensitive to the trajectory form but not to the underlying dynamics of an event. Infants showed a significant dishabituation when the playback direction of point-light events was reversed, indicating that they were sensitive to the switch in the form of the trajectory. On the other hand, no differences were found in the number of trials to reach habituation criterion between forward and reversed events. The authors argue that this reflects a lack of sensitivity for naturally occurring dynamic causes. Pavlova, Krägeloh-Mann, Birbaumer, and Sokolov (2002) examined the perception of point-light displays when the animation of a dog was played backward. The authors observed what they called an “apparent-facing effect”: adults as well as children between 5 and 7 years often misjudged the direction in which a point-light depiction of the dog was facing when the display was reversed. Instead of the correct recognition of the dog under normal conditions, reversed displays were often interpreted as consisting of two or more objects performing a variety of actions. Apparently, the visual system
lacks sensitivity for reverse transformations of point-light displays. Moreover, the percepts were adapted to familiar structures, similar to the findings of Bülthoff et al. (1998).

Jokisch and Troje (2003) investigated whether observers are sensitive to the constant force of earth’s gravity, which defines a relation between temporal and spatial parameters. More precisely, an efficient stride frequency can be predicted by an inverse quadratic relationship, with smaller animals employing a higher frequency. In two experiments, observers had to adjust the size of a walking dog. Displays with high stride frequencies were perceived to be of smaller size than animals with low frequencies, and the correlation between estimated size and stride frequency very closely resembled the predicted inverse quadratic relationship. It was concluded that the motion pattern of an animate being contains sufficient cues to retrieve its size and therefore to scale the visual environment.

Taken together, the findings suggest that the human visual system develops an early sensitivity for dynamic causes of events. The kinematics of a point-light display can be sufficient to convey kinetic properties such as masses or muscular effort, which are in turn picked up and interpreted by human observers. On the other hand, some results indicate that a temporal reversion of a point-light display is not as easily perceived as other deviations from natural laws.

The detection of dynamic causes of an event requires the separation of the underlying structure of the agent from the dynamic changes this structure undergoes. How can these two types of information be segregated, to put them to a closer analysis?

2.5 Structural and kinematic information

In a study dealing with dynamic face perception, Pittenger and Shaw (1975) noted that the perception of any event has two components. One component is the detection of “structural invariants”, that is, invariant information specifying the structure to which movement is applied. Concerning gait perception, static
point-light displays appear as unorganized collections of dots, which are not meaningfully related to each other, but as soon as the displays start to move in a coherent fashion, the structure of the object becomes apparent ("form-from-motion"). These are the physical relations that remain constant throughout the event. The other component identified by Pittenger and Shaw (1975) was the perception of the "transformational invariant", that is, the information about the motion itself, or the kinematic description of how the physical structure changes over time. These two components are similar to what Köhler termed the "topographical" and "dynamic" determinants of the fate of a physical system (1947, p. 107).

Troje (2002a, 2002b) proposed a method to decompose human walking data into structural and kinematic information which is based on either principal component analysis (Troje, 2002a) or a discrete Fourier analysis (Troje, 2002b). An individual’s walking pattern can be represented very accurately by the average posture of a walker including his or her size, the first few harmonics, which represent the local motions of the point-lights, and the walker’s gait frequency. In this way it is possible to map the walking data of an individual into a morphable multidimensional space which in turn provides a basis for the application of linear pattern recognition methods. Furthermore, the Fourier representation not only provides a decomposition into structural and kinematic components, but also allows to separately manipulate attributes such as size, shape, fundamental frequency as well as amplitude and phase spectra of the different Fourier components. This method was used in the studies of the current work to examine the differential influence of several parameters on the identification of individual persons from biological motion.

The perceptual significance of structural and kinematic properties of point-light displays was subject to several studies. Mather and Murdoch (1994) examined whether observers relied more on structural or on kinematic cues to identify the gender of a point-light walker. They synthesized different types of walkers by distinguishing male or female shoulder-to-hip ratio as a structural cue and lateral movements of the upper body as a kinematic cue. By combining these two factors, the authors created different types of walkers with corresponding or contradicting information about the walker’s gender. They
showed that observers relied more on the kinematic than on the structural cue, but only if the walkers were shown from a frontal view, and not from a half-profile view. The latter resulted in very poor gender identification, although much of the relevant information was still accessible. It seemed that in the half-profile view there was no cue in the displays which helped the subjects in their decisions between a male or female walker.

Using motion capture data from 40 walkers, Troje (2002a) synthesized point-light displays by replacing individual kinematic or structural information of the walkers by the mean values of the database. These stimuli were shown to observers who had to classify the gender of the walkers from three different viewing angles (frontal, half-profile, or profile view). Gender classification was best in the frontal view and diminished with an increase in viewing angle. Moreover, performance was significantly better when the displays contained kinematic information than when only structural information was given. These results, as well as those of Mather and Murdoch (1994) indicate that the gender of a point-light walker is expressed primarily in the kinematic cues of a display. Furthermore, a frontal view is most informative for this kind of task, and even a half-profile view leads to a significant impairment of gender discrimination.

### 2.6 The role of the viewpoint

The effect of the viewpoint from which a point-light walker is seen has not received much attention in the literature. For the recognition of static objects, this has been an important issue for a long time. The particularities of our ability as well as our failures to recognize an object from novel viewpoints have provided ample insight into the mental representations and perceptual mechanisms of object recognition. Whereas the discussion about whether visual representations of objects are based on viewpoint dependent two-dimensional projections or on viewpoint independent structural descriptions was originally very controversial (Biederman & Gerhardstein, 1995; Tarr & Bülthoff, 1995), the more recent literature has come up with a more refined view showing that - depending on the particular object class, level of object-specificity, or on
particular tasks - both image-based as well as structural representations can play a role (Foster & Gilson, 2002).

For biological motion, there exist only few studies which have systematically investigated viewpoint dependence. Most of them have been conducted in the context of gender classification. Apparently, the frontal view reveals more information about a walker’s sex than the sagittal view or other views between these extremes (Mather & Murdoch, 1994; Troje, 2002a), as has been discussed above. Verfaillie (1993) studied the effects of several attributes of point-light displays with a priming paradigm and found that only an identical orientation of priming and primed stimulus had a beneficial effect on the reaction time in a recognition task. No effects were found for the walking direction, the starting phase inside the walk cycle, or for an off-joint positioning of the markers on the body of the walker.

In contrast to the dominant role of the frontal view in the perception of biological motion, this effect was not found in studies about face recognition under varying viewpoints (e.g., Bruce, Valentine, & Baddeley, 1987; Krouse, 1981; Logie, Baddeley, & Woodhead, 1987; Troje & Bülthoff, 1996). In general, these studies show an advantage of the half-profile view, whereas for biological motion there seems to be an advantage of the frontal view, at least in tasks that aim at gender classification. Thus, although it is difficult to compare studies about face perception and biological motion because of the distinct objects of interest, there seems to be a general difference between the effects of varying viewpoints in these two domains. Therefore, one of the aims in this work is to clarify the role of the viewpoint for the perception of biological motion.

2.7 Person identification from biological motion

2.7.1 The recognition of gait patterns

In contrast to the well documented, but nevertheless still impressive performance of the visual system to organize isolated point-lights into a coherent figure, the principles underlying the retrieval of specific information
from biological motion are only poorly understood. This is in particular the case for judging person identity from point-light displays. In a natural setting, recognition of known people is confounded by familiarity cues such as face, clothing, or hairstyle, and by the probability of meeting someone at a given place and time. Thus, the usage of point-light displays allows the investigation of the role of purely structural or kinematic cues for person identification. Nevertheless, only few studies (Beardsworth & Buckner, 1981; Cutting & Kozlowski, 1977; Jokisch, Daum, & Troje, submitted; Stevenage, Nixon, & Vince, 1999) engaged in the assessment of how people achieve to recognize individual motion patterns from these point-lights.

The first experiment on person identification from biological motion was conducted by Cutting and Kozlowski (1977). The authors applied the technique suggested by Johansson (1973) to record the gait patterns of a group of six persons who were familiar to each other. The actors were filmed from a profile view, striding back and forth in front of the camera. The same actors served as observers in a subsequent test session with free response trials. Thus, they were familiar with the individuals displayed but not with their point-light depictions. The task was to repeatedly assign the point-light displays to the individual actors, including oneself. Observers started with a relatively low performance, but even though no feedback was given, they improved to a level of correct responses which was well above chance performance. Self-recognition was not significantly better than the recognition of others (43% vs. 36%). When asked about their strategies, observers mentioned characteristics such as speed, rhythm, amount of arm swing, and length of steps helping them to identify the walkers. Whereas Johansson (1973, 1976) had shown that a dynamic array of point-lights is sufficient to detect the presence of a human structure, this study was the first to demonstrate that the array contains sufficient information to identify individual motions.

Beardsworth and Buckner (1981) remarked that Cutting and Kozlowski (1977) failed to control the number of guesses in favor of different individuals. Hence, the participants might have been biased in choosing a particular answer, particularly in favor of or against oneself. Therefore they replicated the experiment of Cutting and Kozlowski (1977) with the same number of subjects,
but constrained the number of each subject’s guesses about himself and the others to the same number as the respective presentations were given. This means that from a total number of 60 walker presentations, the number of each subject’s guesses about himself was constrained to 10, whereas in the remaining 50 presentations the subject had to choose a different person. Moreover, they used longer display durations and a slightly more complex motion pattern. This time a significant difference between self- (58%) and other-recognition (32%) was found. It is not clear if this difference, compared to the study of Cutting and Kozlowski (1977), is based on the constrained number of answers, or on the broader pattern of motions.

Stevenage et al. (1999) videotaped six walkers whose identifying characteristics such as face, hair or detailed body shape were masked, under different display conditions (daylight, dusk, or point-light). The observers in this study were not familiar with the actors, and they had to learn to discriminate the walkers via feedback until a criterion level of learning was reached. Female walkers could be identified more easily by the observers than male walkers, but neither the gender of the perceiver nor the display condition had a significant effect on identification performance. The latter finding was counted as strong evidence that observers used gait-related rather than body-related cues to solve the identification task, because the different amounts of information about general shape in the three display conditions had no apparent effect on recognition performance. Whereas peer groups of actors and observers were acquired in the two studies mentioned previously, Stevenage et al. (1999) were the first to show that it is possible to learn to discriminate between individual gait patterns of people who are not known to the observers beforehand.

Jokisch et al. (submitted) used the motion capture technique mentioned previously to study the recognition of self- and other-generated gait patterns under varying viewpoints. Twelve subjects, whose point-light depictions were normalized with respect to their overall sizes, served as actors and observers. The displays were shown from either a profile view, a frontal view, or a half profile view, the latter being specified as rotated by 30° against the frontal view. It was found that the viewpoint only affected the recognition of other persons. Whereas the identification of one’s own gait pattern was equal under all three
viewing angles, recognition of other persons was significantly impaired from a profile view. This is in accordance with the results of Beardsworth and Buckner (1981), who only used profile views and found a similar discrepancy between self- and other-generated movements.

In summary, it has clearly been shown that observers are able to recognize individual persons by their motion patterns. Those can be familiar persons, but it is also possible to learn to differentiate previously unknown people by the way they walk. Some results indicate that there might be a viewpoint effect in recognizing self- and other-generated movements, but further and more detailed research is needed to support these findings. Moreover, it is still unclear whether there are specific factors that are used to solve the identification task, especially concerning parameters that are intrinsic to the walkers. The study of Jokisch et al. (submitted) indicates that retaining the individual size of the walkers is not necessary for the recognition, because the point-light stimuli in this study were normalized with respect to their overall sizes. On the other hand, there is no baseline which indicates whether this size-normalization impairs walker recognition. Thus, the first two studies of the present work investigate the effects of normalizing specific walker-parameters like size, body shape, gait frequency, or kinematic details of the gait pattern.

Shiffrar and colleagues (Jacobs, Pinto, & Shiffrar, 2004; Loula, Prasad, Harber, & Shiffrar, 2005) examined the effects of visual and motor experience on the recognition of biological motion. Jacobs et al. (2004) found that visual experience as well as motoric feasibility can lead to enhanced recognition of a walker, but that the effect depended on the task at hand. The authors conducted a series of experiments in which they examined the sensitivity to point-light walkers across a variety of tasks. Two models served as actors who walked on a treadmill at different speeds, either with a normal gait pattern or with unusually long or short steps. Observers saw point-light versions of these gaits which were played back either forward or reverse, the latter leading to a physically impossible walking pattern. In a task of speed-discrimination, a higher sensitivity was found for possible walking patterns than for impossible ones, indicating a significant effect of motoric feasibility. When the observers were separated into groups of different familiarity levels with the models, this had only
effects on tasks of walker-identification and walker-discrimination: highly familiar observers performed better than observers who were hardly familiar with the models, but only when normal gait patterns were shown. Moreover, observers who were familiar with the actors did not show an increased performance in a speed-discrimination task. Visual experience only seems to play a role in tasks that have a social relevance for the observers. In contrast to these specific results for visual experience, a general finding of Jacobs et al. (2004) was that reversing the animations of the point-lights impaired the performance on all tasks. The authors suggest that this result reflects the effect of motoric feasibility. On the other hand, this cannot be clearly separated from visual experience, because human observers are also more likely to have more experience in watching other humans moving forward than performing a backward directed motion.

Loula et al. (2005) found more direct evidence for the influence of motor experience on the perception of point-light displays. They recorded a variety of motion patterns (e.g., walking, running, dancing, jumping) from pairs of persons who were known to each other and used these recordings to create stimuli of ‘friends’ and ‘strangers’. In a subsequent identification task, the same persons had to indicate whether point-light versions of the motions depicted themselves, their respective friends, or a stranger. Whereas recognition of one’s own movements was best, identification of the friends, but not of the strangers, was also above chance level. These results were confirmed in a subsequent actor discrimination task, to control for a possible bias in naming self or others. The relative sizes of the effects suggest that motor experience contributes the main part to the recognition of the displays, but that visual experience also plays a significant role. The best recognition performance was found for complex motion patterns like dancing, but only low performances for walking and running movements. This is a possible explanation for the differences in the results between the studies of Cutting and Kozlowski (1977) and Beardsworth and Buckner (1981). The latter authors reported an enhanced recognition of self-generated motions, but they had also used a slightly more complex motion pattern. Cutting and Kozlowski (1977), on the other hand, had only recorded the walks of their participants.
The studies of Shiffrar and colleagues indicate a significant effect of experience on the recognition of a person performing an action. The results of Loula et al. (2005) suggest that motor experience plays a dominant role in the identification of the agent, but that the effect of visual experience is not negligible, either. The findings of Jacobs et al. (2004) further underscore this assumption, because there was a generally higher sensitivity for possible motion patterns than for impossible ones. On the other hand, the effects of motor and visual experience were not clearly separated, because a reverse transformation of point-light walkers results in gait patterns which are biomechanically impossible as well as hardly observed in our everyday environment.

In the third study of the current work, visual and motor experience were segregated by recording actors who walked on a treadmill either forward or backward, and by showing these gait patterns either in the normal way or as a reversed display. The enhanced recognition of one’s own gait pattern was only shown in explicit identification tasks, that is, the observers had to identify the actor or to discriminate between self- and other-generated actions. In the present study, it was examined whether visual and motor familiarity with a motion pattern also affects tasks that are not explicitly related to person identification.

2.7.2 Identification studies in other domains

The identification of self- and other-generated actions is not limited to the domain of whole-body motion. An extensive amount of research has been conducted on the identification of speech (e.g., Fellowes, Remez, & Rubin, 1997; Lachs & Pisoni, 2004; Remez, Fellowes, & Rubin, 1997; Sheffert, Pisoni, Fellowes, & Remez, 2002). Moreover, Knoblich and colleagues (Flach, Knoblich, & Prinz, 2004; Knoblich & Flach, 2001; Knoblich & Prinz, 2001; Repp & Knoblich, 2004) investigated the role of the agent in a variety of domains, including the recognition of one’s own clapping, drawing familiar and unfamiliar
characters, the intonation of short musical excerpts, and the landing position of thrown darts.

In speech perception, several studies indicate that talkers are identifiable from their linguistic characteristics, and that experience with a talker affects the linguistic analysis of speech (see Sheffert et al. 2002, for a review). Sheffert et al. (2002) showed that listeners can learn to identify previously unknown talkers solely from phonetic attributes preserved in the individual sinewave signals of training sentences. Moreover, this knowledge generalized to new sentences presented as natural speech or as sinewave versions. In contrast, training with natural speech did not generalize well to test sentences with sinewave utterances. Listeners were also able to recognize talkers on the basis of qualitative attributes, from incomprehensible reversed speech samples.

Flach et al. (2004) investigated whether listeners could recognize one’s own clapping. It was demonstrated that acoustic recordings of clapping enabled the participants to discriminate between self- and other-generated sounds, even when the claps produced uniform tones. The temporal variation was found to be an important factor for recognition. Knoblich and Prinz (2001) showed that subjects could identify self-generated actions from kinematic displays of their own drawings. The subjects had to draw familiar and unfamiliar characters without receiving visual feedback. In a later recognition experiment, the drawers were able to discriminate the kinematics of their own drawings from the ones of other subjects, even if the drawing actions were scaled so that they matched in overall size and duration. Performance dropped to chance level only if the drawings were shown with a constant drawing velocity, so that the temporal structure of the drawings was broken. A further study (Knoblich, Seigerschmidt, Flach, & Prinz, 2002) demonstrated that observers were more accurate in predicting subsequent handwriting trajectories when these were generated by themselves, compared to the trajectories of others’ handwritings.

Repp and Knoblich (2004) examined the factors that determine if pianists can recognize whether they played a musical excerpt themselves or whether it was played by another person. Experienced pianists had to play short parts from pieces of classical piano literature, either with or without auditory feedback over headphones. In a later test session they had to listen to each recorded
excerpt and indicate on a scale how sure they were that the pieces were played by themselves. Highest ratings were always given to the recordings that were indeed played by the pianists themselves, even if these were normalized with respect to overall tempo and dynamics. This result was independent of auditory feedback. Knoblich and Flach (2001) showed that observers could predict the landing positions of thrown darts more accurately when they watched their own throwing movements on a video recording than when they watched the recordings of other persons. This was the case when the observers saw the full video, only the body, or only the arms of the actors, but only in the second half of all trials. The authors suggest that the observers initially relied on a third-person perspective when they judged the throws, and only after a while switched to a first-person perspective when they saw their own actions.

Taken together, the agent of an event can be identified from a number of domains. Whole-body movements as well as single trajectories, for example, those of handwritings, are sufficient to recognize the person who generated an action, and there seems to be an enhanced sensitivity for motions that have been conducted by oneself, either from visual or from auditory stimuli.

2.7.3 The understanding of actions

The “common coding” approach, as advanced by Prinz (1997) focuses on the specific effects that follow self- and other-generated actions. The main assumption is that actions, whether they are executed or perceived, are coded in terms of the effects they are supposed to generate. Thus, an extraneous effect is mutually represented by its perception as well as by the action plan generating this effect. In the case of walking patterns, one would assume that the perception of a point-light walker leads to a parallel activation of those parts of the motor system that are involved in the creation of the gait pattern. When an observer views a self-generated motion pattern, there is a closer correspondence between the effects that he perceives in the displays and the effects that would have followed a self-generated action plan, than this would be the case for movements generated by different agents. Although human bodies
are based on the same general shape, they differ in their specific anatomical constraints as well as in the way they have learned to perform characteristic actions.

Research in neuroscience suggests a common coding at the neuronal level. Gallese, Fadiga, Fogassi, and Rizzolatti (1996) discovered that neurons in the macaque's premotor cortex discharge not only when the monkey is performing a goal-directed action (grasping) but also when the monkey is watching someone executing the same action. Thus, mirror neurons represent particular actions, independent of the performing agent. These results were confirmed in a number of brain imaging studies, for example, by Rizzolatti, Fadiga, Gallese, and Fogassi (1996). Ferrari, Gallese, Rizzolatti, and Fogassi (2003) found similar mirror activation in monkeys when they observed another individual performing mouth movements. Motor-related areas were also found to be active during point-light perception of visible speech (Santi, Servos, Vatikiotis-Bateson, Kuratate, & Munhall, 2003).

Saygin, Wilson, Hagler, Bates, and Sereno (2004) extended the findings by showing that there is even a robust response to biological motion in the human inferior frontal and premotor cortex (PM). These areas, previously known to respond to action observation, showed selective activation, compared to static or scrambled point-light stimuli. Thus, even degraded versions of actions, which are solely characterized by motion cues, are sufficient to activate this action observation system. The authors suggest that the motor system of the observer is recruited to replenish these incomplete stimuli.

In a recent study, Sakreida, Schubotz, Wolfensteller, and von Cramon (2005) examined differential PM activation under the perception of distal (fingers, mouth), proximal (knee, ankle, elbow, wrist), or axial (trunk, shoulder) intransitive (i.e., non-goal-directed) movements. Whereas distal motions produced most significant activation in ventrolateral PM areas, viewing proximal motions lead to an enhanced activation of the dorsolateral PM. Axial motion elicited the strongest activation in the medial PM (supplementary motor area). Thus, these findings support an observable but limited somatotopy in frontal motor cortices.
Recent research indicates that visual self-recognition is mainly processed in the right hemisphere, specifically in the right frontal lobe (e.g., Keenan, Freund, Hamilton, Ganis, & Pascual-Leone, 2000; Platek, Keenan, Gallup, & Mohamed, 2004; Sugiura, Watanabe, Maeda, Matsue, Fukuda, & Kawashima, 2005), although these findings only relate to the recognition of one’s own face. Thus, it is not clear whether the perception of self-generated actions, either full-body videos or point-light depictions, follow the same pathway.

2.8 Summary

Human observers are experts in recognizing biological motion. At least from the age of 3 months on, infants exhibit sensitivity for coherent point-light displays, and at the age of 5 months they seem to have developed a global concept of a walking human figure. Because it has a high ecological importance, even small anomalies in natural motion patterns are detected by the visual system. The inversion of a point-light walker severely impairs its analysis, even up to a complete failure to recognize it as a walking figure. The questions about the exact processes which underlie the general perception of biological motion are still unsolved. Neither bottom-up analyses nor top-down processing can solely account for the findings which have been presented in this chapter. The results obviously depend on the task at hand, and attention as well as prior expectations about what we are going to see play a role in the recognition of the stimuli.

Findings from studies about person identification clearly show that human observers are able to discriminate individual persons shown as point-light walkers. Apparently, pictorial information like face, clothing, colour of hair or skin, or other visual details about the person are not necessary to successfully solve this identification task. The persons to be recognized can be unknown actors as well as walkers who are personally known to the observers, or even include themselves. Nevertheless, it is unknown which factors of the point-light depictions significantly contribute to the recognition. For example, none of the previous studies applied methods to separate structural from
kinematic information to analyze these two factors individually. Furthermore, it is unclear whether there exists an enhanced sensitivity for the perception of self-generated gait patterns and whether there are conditions that advocate this sensitivity.

2.9 Objectives of the current work

The main objective of this work is to reveal the factors underlying the identification of individual persons from biological motion. Across a variety of domains, the identity of an agent performing an action can be separated from the action itself or from its meaning. The meaning of an action is an abstract representation, independent of the specific conveyance through an individual person. The most common example of this separation is speech perception, in which identical spoken words denote the same contents, detached from a specific person. On the other hand, it is possible to identify individual persons from very limited phonetic properties. Furthermore, listeners are able to learn to discriminate talkers from these reduced properties, and to generalize this knowledge to natural language.

In the present work it was examined whether similar principles can be applied to person recognition from biological motion. The first study was designed to investigate which parts of the overall information contained in biological motion have an effect on person identification. Observers learned to discriminate point-light walkers whose depictions were manipulated along two different lines. First, the walkers were shown from three different viewpoints. Second, parts of the individual, diagnostic information such as size, shape, and step frequency were gradually replaced with averaged values, rendering the respective part of the overall information useless for the identification task. In addition to measuring learning curves in the reinforced training sessions, non-reinforced test sessions were introduced at different stages of the learning process. Most of them presented the walkers from the same viewpoint as during training. Only at the very end of the experiment, a further test session was added in which the walkers were presented from novel viewpoints. A significant
generalization of identification performance would indicate an abstract representation of the walking patterns, which is not solely based on the trajectories of the markers as seen from a specific viewpoint.

Whereas training started with the original, unaltered versions of the point-light stimuli in Study 1, the second study applied only walkers who were normalized with respect to size, shape, and gait frequency from the first session on. This design was used to examine whether structural cues are necessary to learn to differentiate individual point-light walkers, or whether observers can rely solely on kinematic information. Moreover, the focus of attention was on the role of specific kinematic information of the walkers, namely on the effect of normalizing individual harmonics of a walking pattern. Whereas the physical space is dominated by the first harmonic, which explains more than 90% of the overall variance of a gait pattern, it is not clear whether this also applies to the psychological representation of the walkers, or whether higher-order harmonics also contain sufficient or necessary information to recognize an individual walker. In a second experiment, the differential relevance of the phase- and amplitude spectra was analyzed. In both experiments, a similar learning paradigm as in Study 1 was used. In the first experiment, the training part was followed by a non-reinforced test session in which the stimuli were normalized with respect to different harmonics, and a final test session in which the viewpoints under which the walkers were shown were additionally varied. The second experiment followed the same design, with the exception that the test stimuli were normalized either for the amplitude or for the phase spectrum of the walker.

The third study examined the effects of visual and motor experience with gait patterns on a number of perceptual tasks. Several studies have shown a close connection between the generation of actions and the subsequent recognition of kinematic patterns of those actions. However, in all of these studies comprising biological motion, observers had to explicitly discriminate between self- and other-generated motions. Under such conditions stored representations of one’s own motion patterns, or “sprites”, as Cavanagh et al. (2001) have called them, are probably activated and compared to the displays which are presented on the screen. Thus, judgement about identity is reached
by a direct comparison between stored and perceived motion patterns. Hence, a number of indirect measures were applied to explore the asserted enhanced sensitivity for self-generated actions in a different context. The walking patterns of ten actors were recorded on a treadmill at different speeds, and the point-light depictions were used for subsequent perceptual tasks with the same actors in the role of the observers. Tasks included the detection of a coherent figure in a scrambled mask, the discrimination of the playback direction of an animation, and the active adjustment of the estimated original gait frequency of the walker. On each task, point-light displays were used which showed either the observers’ own motion patterns, the movements of a familiar person or those of strangers. A robust elevated sensitivity for self-generated actions should manifest itself even in these indirect perceptual tasks. Furthermore, a final experiment tested identification performance under varying viewpoints and display conditions.
Chapter 3: General Methods

3. General Methods

3.1 Data collection

The gait patterns of all models were recorded using a motion capture system (Vicon; Oxford Metrics, Oxford, UK). In a motion capture session, a set of 41 retroreflective markers was attached to the body of the subject (see Fig. 3.1). The models wore gymnastics suits and shoes, and most of the markers were attached directly to the skin or to the suit. The markers for the head, the wrists, and the ankles were positioned using elastic bands, whereas the markers on the feet were attached to the shoes.

The system’s cameras emit a red light that is reflected by the markers and gathered again by the cameras. Thus, the cameras do not record images of the subject himself, but they only trace the trajectories of the markers. The positions of the markers are tracked with a spatial accuracy of 1 mm and a temporal resolution of 120 Hz.

![Figure 3.1: Placement of the 41 markers on the body of the actor.](image-url)
The cameras were mounted at the sides and at the ceiling of a room of 11 m length and 7.5 m width, in order to allow for a maximally expanded capture volume, that is, the range in which all markers on the subject’s body can be traced accurately. To ensure a 3-dimensional tracking, each marker has to be detected by at least two cameras at all points inside the capture volume.

Recordings of freely walking subjects were conducted with the participant walking across this volume, which was 6 m in length, allowing for approximately four gait cycles. When the subjects were recorded walking on a treadmill, this was placed at the centre of the capture volume. Recordings of walking patterns on the treadmill comprised about 10-15 gait cycles.

The individual images provided by the cameras were collected by a data station that computed the positions of the individual markers on each frame. Afterwards, the 3-dimensional trajectories of the markers were reconstructed from these data using commercially available software (Workstation, Oxford Metrics). Possible gaps in the individual trajectories were filled using either an interpolation technique, or by adopting the respective trajectories of adjacent markers which were connected by a rigid segment. For example, gaps in the trajectory of a knee marker could be filled by adopting the trajectory of the upper leg marker throughout this time span. Following this reconstruction, biomechanical modelling (Bodybuilder, Oxford Metrics) was used to compute 15 “virtual” markers, that is, the positions of the major joints and other relevant structures inside the body were estimated from the original data. These 15 markers represent the centre of the head, the shoulders, the elbows, the wrists, the centre of the pelvis, the sternum, the hips, the knees, and the ankles. Finally, the data were imported into Matlab, using custom designed software. The common translatory component of each marker was removed from the data, so that the walker appeared to walk on a treadmill in all cases.

3.2 Decomposition of the data

The final data eventually represented a time series of postures with 120 entries per second, each consisting of a 45-dimensional vector which contained the 3D cartesian positions of the 15 virtual markers. This time series was
decomposed into a low-order Fourier expansion for each walker. Only complete gait cycles of a walking pattern were used for this expansion. The time invariant part of this decomposition contained structural, anthropometric data (the average posture of a walker over a complete gait cycle), whereas the time dependent terms contained information about the kinematics of the single markers (Troje, 2002b). Thus, each gait pattern was represented as a matrix of structural and kinematic parameters for all 15 markers in three dimensions. Furthermore, the period of the gait pattern, that is, the time needed to complete a gait cycle, was computed. Moreover, the overall size of the walker, compared to a prototype, was calculated. Each parameter of the decomposed data was divided by the individual size, and a single value for the size was added to the walker representation. Finally, the gait pattern was phase-shifted, so that each walker’s gait started at the same point of the gait cycle. From this matrix it was possible to reconstruct the posture of a walker at each point of time by adding a weighted sum of the Fourier components, based on the position on the gait cycle, to the average posture and by multiplying the result with the relative size of the walker.

After applying these methods, the resulting representations of the individual walkers were morphable, that is, linear combinations of existing walkers resulted in naturally looking walking patterns. In particular, averaging the data across a number of walkers resulted in a walking pattern that appeared to be as real as any of the constituting walkers. This was also the case for discrete parts of walking patterns. For example, it was possible to normalize a group of walkers with respect to the structural information by replacing all individuals’ time invariant terms with the average of these terms across the whole group. This resulted in point-light walkers which were based on the same average body shape but which still varied in terms of their kinematic patterns.

This method reflected the distal-to-proximal approach, as proposed by Runeson (1994). The data material was based on raw motion capture data, that is, on the trajectories of the markers on the actor’s body. The Fourier transformation very closely resembled the original motion pattern captured by the system. The first two harmonics captured on average more than 97% of the overall variance in a gait pattern, therefore providing an accurate representation of the distal actor. Based on these data it was possible to vary discrete
parameters of the walking patterns to examine how they differentially affected the recognition of the walkers.

A major advantage of the motion capture technique, compared to video-based methods, is the independence of specific viewpoints. The trajectories of all markers are recorded by nine cameras that are positioned around the actor, leading to a complete three-dimensional representation. Thus, when a point-light walker is shown on the screen, this stimulus is based on the projection of a three-dimensional gait pattern onto the fronto-parallel plane. Moreover, the stimulus can be rotated around all axes to examine the potential role of the viewpoint on the perception of biological motion.

For the first study, as well as for the first experiment of Study 2, a self-developed program for Linux was used to display the stimuli on the screen. This program used a set of previously computed gait patterns for each walker that were stored in ASCII-files. Each file started at a different point in the step cycle of the walker to create different configurations of the point-lights at the start and at the end of each animation. For the second experiment of Study 2, as well as for all experiments Study 3, the displays were shown using Matlab and the Psychophysics Toolbox (Brainard, 1997). In this case, the matrix-representation of a walker was used to dynamically generate the position of each marker at every frame that is shown on the screen. A randomly determined starting point in the gait cycle was accomplished by referring to the internal clock of the computer.
4. Study 1

4.1 Introduction

This study was designed to investigate the differential role of structural and kinematic cues for person identification from biological motion. The observers were trained to name seven male point-light walkers over several reinforced training sessions. In the course of the training the parameters of size, shape, and gait frequency were normalized and thus rendered uninformative for the task at hand. The experiment was conducted over five consecutive days. At the end of each day, a single non-reinforced test session was implemented in which the identification of the walkers under each normalization condition was tested. Whereas each observer saw the stimuli only from one viewing angle, this viewpoint was varied in the final test session to examine perceptual generalization of the walking patterns over multiple views.

4.2 Methods

Participants:
Eighteen observers (12 women, 6 men, age 21-40 years), all students or employees at the Ruhr-University in Bochum, participated in the experiment. All subjects were naïve to the purpose of the study and had no previous experience with point-light displays. Students received course credit for their participation. All observers had normal or corrected-to-normal vision.

Stimuli:
Seven male individuals served as walking-models. All of them were students or staff at Ruhr-University, between 25 and 35 years of age. They were asked to walk on a treadmill at a self-adjusted speed for at least 10 minutes until they felt perfectly comfortable with the setup. Not telling them when data acquisition was actually turned on, a walking sequence of 10 sec was recorded. The motion capture system and the techniques described in Chapter 3 were used to record
the gait patterns of the actors and to compute the positions of the respective 15 virtual markers.

Average posture, relative size and walking frequency of each walker were used to create a number of modifications from the original data. Five different conditions were used in the experiment:

1. In the veridical condition (VR), the original data of the walkers as described above were shown to the observers.
2. For the size-normalized condition (SI) each walker’s complete three-dimensional data were divided by his relative size. To compute relative size a number of steps was involved. First each walker’s average posture was computed by averaging all frames across one full gait cycle. An overall average posture was then computed from the seven individual average postures. Now each individual's relative size was determined by the slope of the regression between this individual's average posture and the overall average posture. The seven relative sizes were: 1.10, 1.03, 0.98, 0.95, 0.89, 0.98, 1.07. After dividing each walker’s data by his relative size every walker now had the same relative size 1, but the relative positions of the point-lights still differed between the walkers, retaining each individual’s shape and kinematics intact.
3. The stimuli for the shape-normalized condition (SH) were derived from the size-normalized data by replacing each walker's individual average posture with the overall average posture. This was achieved by simply subtracting the individual average posture from each frame (this time computed from the size-normalized data) and then adding the overall average posture to it. The resulting walkers now all had the same average posture. Before shape-normalization the mean variance of the location across the 15 dots was $612 \text{ mm}^2$.
4. Frequency-normalized stimuli (FR) were obtained directly from the veridical data. Using Fourier analysis, each walker’s step frequency was determined. Frequencies for the seven walkers were: 1.42, 1.29, 1.04, 1.11, 1.25, 1.19, 1.13. Then the data were re-sampled such that they all walked with a frequency of 1.2 Hz, which is the average across the
seven walkers. The walkers still contained their individual sizes and shapes.

5. For the last condition (AL) the same procedure to normalize for frequency was used, but this time with the shape-normalized stimuli. The AL-data thus lack diagnostic information about size, shape, and frequency, but still contain differences in amplitudes and relative phases of the motions of the single dots.

Note that the four conditions VR, SI, SH, and AL are formed by gradually omitting information about size, shape, and gait frequency, respectively.

The point-light stimuli were shown to observers on a 19" CRT monitor (1240 x 1024 pixels) as white dots on a black background from one of three possible viewing angles: 0° (Frontal View, FV), 30° (Half-profile View, HV) or 90° (Profile View, PV), as can be seen in Figure 4.1.

The size of the walkers on the screen was about 6 cm, corresponding to a visual angle of 3.4° at the viewing distance of 1 m. No possible occlusions were modeled, that is, all 15 markers were visible in every frame and from each
viewpoint. Each display lasted for 3 s, during which about 2.5 full gait cycles were shown.

*Design and Procedure:* 
Participants were randomly assigned to one of three groups according to the viewpoint from which the walkers were shown (factor VIEW with levels FV, HV, and PV). Each group contained six observers. Viewing angles remained constant for each observer except for the final test session. For the HV and the PV groups, half of the participants saw the walkers facing to the left side and half saw them facing to the right side. This design was used to counterbalance possible effects due to hemispheric asymmetries, without intending to systematically investigate these effects.

Every experimental session consisted of a specified number of trials. A trial began with the 3 s stimulus display, followed by the presentation of seven buttons, labeled with fictitious names of the walkers. Participants had to use the computer mouse to click on one of them in order to name the walker. The next display followed after an inter-trial-interval of 1 s.

The experiment consisted of a series of 9 training sessions, which were interleaved with 4 type I test sessions (see below). At the end of the experiment, a single type II test session was added.

In each training session, one of four different types of walking data was used (factor NORM with levels VR, SI, SH, and AL), thus gradually decreasing the amount of available diagnostic information. Training sessions consisted of 140 trials, which were subdivided into 10 blocks of 14 trials. In every block, each of the walkers was shown twice in randomized order. If the participants clicked on an incorrect name, the button with the correct name changed its color for 1 s, thus providing feedback. If the answer was correct, nothing else happened before proceeding to the next trial.

In the type I test sessions all five types of stimuli were used (VR, SI, SH, FR, AL). The sessions consisted of 140 trials (7 walkers x 5 conditions x 4 repetitions) in randomized order. The viewpoint in the type I test sessions was the same as the one that was used in the training sessions for each observer. No feedback was given to the participants.
In the type II test session only the all-normalized data (AL) of each walker were used, but the viewpoint of the displays was varied: the point-light displays were shown from all three different viewpoints. The session consisted of 147 trials (7 walkers x 3 viewpoints x 7 repetitions) in randomized order, and no feedback was given to the participants.

The experiment was run over five consecutive days. On the first four days, there were two training sessions followed by a type I test session. The training sessions on the first day consisted of only veridical stimuli (VR). Each of the following three days started with one session of re-training with the same stimuli as on the previous day, followed by a second training session with gradually normalized stimuli. On the second day, size-normalization (SI) was introduced, on the third day shape-normalization (SH), and on the fourth day frequency-normalization (AL) was added. The fifth day started with one session of re-training (AL) followed by a single type II test session (Fig. 4.2).

![Figure 4.2: Learning curves of the training sessions. Vertical lines represent the introduction of a new kind of stimulus data. Horizontal arrow indicates chance level. Solid vertical arrows denote the insertion of type I test sessions. The dashed vertical arrow indicates the insertion of the type II test session. The ticks represent the beginning of a new session.](image)
4.3 Results

No significant differences were found between male and female participants for any of the dependent variables, neither during training nor in the test sessions, so the results presented here are pooled over both groups.

Training sessions:
Starting from about chance level, all three groups quickly improved their performances in the identification of the walkers (Fig. 4.2). At the end of the third training session, the FV-, HV-, and PV-groups reached correct response levels of 88%, 70%, and 75%, respectively (average across the last 3 blocks of the session), and after four sessions performance saturated at a level of about 80-90%.

Over the complete series of training sessions, there was a significant main effect of factor VIEW \([F(2,15) = 7.55, \; p = .005]\). The mean values of correct identification were 80% (FV), 67% (HV), and 72% (PV). Bonferroni-corrected Post-Hoc-Tests \((p < .01)\) revealed that the FV-group performed significantly better than the HV-group. No differences were found between the FV- and the PV-groups, or between the HV- and the PV-groups, respectively.

For subsequent data analysis, only the identification performances of the last three blocks before and the first three blocks after changing the level of the NORM condition were used. For example, the SI-level of the NORM condition contained the comparison of the results of the last three blocks with veridical motion to the first three blocks with size-normalized stimuli. By doing this, all relevant data for comparisons were included, and data which could distort further analyses, for example, the low performance values at the beginning of the experiment, were eliminated.

A 3-way repeated-measures ANOVA with CHANGE (with levels Before and After) and NORM (3 levels: SI, SH, AL) as within-subjects factors and VIEW (3 levels) as a between-subjects factor was conducted. Significant effects for all three factors were found: CHANGE \([F(1,15) = 15.00, \; p < .01]\), NORM \([F(2,30) = 14.00, \; p < .01]\), and VIEW \([F(2,15) = 5.21, \; p < .02]\). There was also a significant interaction between CHANGE and NORM \([F(2,30) = 17.12, \; p < .01]\).
as well as a 3-way interaction \[F(4,30) = 2.75, p < .05\]. Post-hoc Bonferroni-tests \((p < .05)\) revealed that on average, identification performance before the introduction of a new dataset (84%) was better than afterwards (77%). Furthermore, the FV-group (87%) performed significantly better than the HV-group (74%), whereas no differences were found between the FV- and the PV-groups (80%), or between the HV- and the PV-groups. The mean performance level before and after the introduction of shape-normalization (84%) was significantly better than the level at the introduction of size-normalization as well as frequency-normalization (78% in both cases). This latter effect is simply due to the fact that performance had not reached saturation by the time the size-normalized versions of the walkers were introduced.

More interesting, however, is the interaction between factors CHANGE and NORM, because it shows the effects of removing diagnostic information from the displays. To further explore this interaction a 2-way, repeated-measures ANOVA was conducted with CHANGE as a within-subjects factor and VIEW as a between-subjects factor, separately for each condition of NORM (Fig. 4.3).

For size-normalization (transition from VR to SI), no significant effect of factor CHANGE \([F(1,15) = 0.95, \text{n.s.}]\) was found. In contrast, this factor had a highly significant effect when normalizing the stimulus for shape (SI to SH) \([F(1,15) = \text{...} \right]
29.44, \( p < .001 \), as well as when removing frequency information (SH to AL)
\([F(1,15) = 28.81, p < .001]\). Significance for factor VIEW was only reached when
changing from SH to AL \([F(1,15) = 6.06, p < .02]\), even though trends confirming
the main effect of VIEW in the 3-way ANOVA were observed in all cases. It
seems like the frontal view is somewhat advantageous compared to the other
two views. No significant interaction between normalization condition and
viewing angle was found in any of the analyses.

**Type I test sessions:**
A repeated-measures, 3-way ANOVA with test-number (NUM, 4 levels) and
type of normalization (NORM, 5 levels) as within-subjects factors and VIEW (3
levels) as between-subjects factor was performed. Identification performances
corresponding to the different levels of the three factors are shown in Figure
4.4. Note that the types of normalization in these test sessions represent the
effects of averaging the single parameters, in contrast to the additive
normalization procedure throughout the training sessions.

Significant main effects were found for every one of the three factors: NUM
\([F(3,45) = 52.84, p < .001]\), NORM \([F(4,60) = 47.77, p < .001]\), and VIEW
\([F(2,15) = 9.32, p = .002]\). Bonferroni-corrected Post-Hoc-Tests \((p < .02)\) were
performed to test for differences between the factor levels. A significant
increase in performance was found between every one of the four test sessions
(53\%, 67\%, 76\%, and 83\% identification rate, respectively), as well as between
most of the normalization conditions. No difference was found between the VR /
SI conditions (76\% / 79\%) and the SH / FR conditions (69\% / 68\%),
respectively. Every other comparison between the conditions reached a
significant level. In addition, the differences between the three viewpoint groups
found in the training sessions were confirmed by the results of the test sessions:
the FV-group (79\%) had a better average rate of identification than the HV-
group (64\%) and even showed a better performance than the PV-group (67\%),
but the HV- and the PV-groups showed no significant differences.
Figure 4.4: Main effects of the three factors used in type I test sessions on identification performance: (A) Number of test session (NUM); (B) Type of normalization (NORM); (C) Training viewpoint (VIEW). Error bars indicate standard error of the mean. Arrows indicate chance level.
A significant interaction between the factors NORM and VIEW \( F(8,60) = 5.01, \ p < .001 \) indicates that different kinds of information are used in the three groups to identify a walking person (Fig. 4.5). The VR-, SI-, and SH-conditions show relatively similar patterns: the frontal view produces the best performance, and the profile view seems to result in a better performance than the half-profile view. This is different in the two conditions which lack information about walking frequency (FR, AL). The lack of frequency seems to particularly affect observers from the PV group. Here, performance drops to the same or even lower values than in the HV group.

There was also a significant 3-way-interaction between the factors NUM, NORM, and VIEW \( F(24,180) = 2.05, \ p = .004 \). It appeared that the increase in performance over the four test sessions became higher with increasing steps of normalization particularly in the PV-group, whereas there were different patterns in the other two groups. In other words: the more information was taken out of the point-light displays, the more time it took particularly for the PV-group to accommodate to the new stimuli.

**Type II test session:**

The walkers shown in the final type II test sessions were normalized according to shape (including size) and walking frequency. In contrast to the type I test...
sessions the walkers were now presented from all three viewing angles. In this way, generalization from one angle to novel viewpoints was tested. A repeated-measures ANOVA was conducted with the training angle (TRAIN, 3 levels) as between-subjects factor and test angle (TEST, 3 levels) as within-subjects factor. Neither factors TRAIN $[F(2,15) = 2.73, \text{n.s.}]$ nor TEST $[F(2,30) = 1.31, \text{n.s.}]$ revealed a significant main effect. However, there was a highly significant interaction between these two factors $[F(4,30) = 19.01, p < .001]$.

As can be seen from Fig. 4.6, the identification performance of the observers was always best if the stimuli were shown from the same viewing angle as in the training sessions. Even though they decreased as the test viewpoint deviated from the training viewpoint, performance levels were still high above chance level in all conditions.

![Figure 4.6](image)

Figure 4.6: Identification performance in the type II test session as a function of Test view and Training view. Error bars indicate standard error of the mean. Arrow indicates chance level.

Further $t$-tests were conducted to examine whether the generalization from one viewpoint to another was significantly better than its counterpart, for example, whether the generalization from the half-profile view to the profile view was better than vice versa. No significant differences were found (all $p$s > .05).
4.4 Discussion

All participants quickly learned to name seven individual male walkers from point-light displays. Starting at chance level, performance saturated at a level of 80-90% correct responses after five training sessions. There is a small but significant frontal view advantage which becomes apparent not only in the learning curve but also from the data obtained from the type I test sessions. Here, the frontal view advantage becomes particularly apparent in conditions in which the observers were deprived of information about frequency.

This pattern is consistent with earlier findings on sex classification from biological motion (Mather & Murdoch, 1994; Troje, 2002a). Important diagnostic features which help discriminating between male and female walks seem to be the rotation of the hip in the fronto-parallel plane, the degree of lateral body sway, the ratio of hip and shoulder width, and the position of the elbows. All these features are expressed in terms of variations in the fronto-parallel plane and are, therefore, best visible from a frontal view. In the present study the task was not sex classification but person identification. Even though it is less clear which kinds of features are most helpful to solve the task it seems that again motion and static cues in the fronto-parallel plane are more diagnostic than from other viewing angles.

The pattern measured here is clearly different from the pattern measured in a variety of face recognition tasks (Bruce et al., 1987; Troje & Bülthoff, 1996) where one generally finds a half-profile view advantage. Because this specific advantage is particularly obvious in experiments investigating the ability to recognize faces from novel viewpoints, it has been interpreted in terms of the fact that faces are approximately bilateral symmetric. Assuming bilateral symmetry, a non-accidental view of an object could be used to recover the three-dimensional affine structure of the object (Troje & Bülthoff, 1998; Ullman & Basri, 1991; Vetter, Poggio, & Bülthoff, 1994), resulting in invariants that would facilitate recognition across viewpoints. In principle, the same argument could hold true for walking as well. Even though each single posture is not bilaterally symmetric, the pattern shows glide reflection, that is, the spatio-temporal symmetry obtained from mirror reversal and subsequent phase-
shifting. It could therefore be used in the same way as proposed in the case of face recognition. In fact, it seems that in the viewpoint generalization experiment (type II test session) the frontal view advantage found in the learning curve and in the other test sessions is cancelled.

The decrease in performance in response to viewpoint changes is very obvious and in general accordance with earlier findings (Verfaillie, 1993). However, the remaining recognition performance is quite substantial. Even under conditions of maximum rotation of the displays between training view and test view (FV to PV and vice versa), recognition performance remains at a level of about 40%, which is still nearly three times higher than chance level (14%). Obviously, the human visual system is very well able to abstract motion seen from a single viewpoint and to generalize it to other viewing angles. This means that there are certain features in human gait patterns which can be recognized without reference to particular viewpoints. Because stimuli in the type II test session were normalized with respect to structural features, the invariants mediating this ability must be contained in the kinematics of the motion itself. Frequency is the only aspect of kinematic information that was looked at in the experiment reported here, and it seems that even though it does affect recognition it certainly does not play a major role. One way to further decompose the remaining kinematic information is to look into the Fourier composition of the walking pattern in order to determine which parts of the power spectrum or the phase spectrum are carrying important information for person identification. These aspects of the gait patterns were examined in the following study.

The learning curve as well as the data from the type I test sessions provide further insight into which features are used for identification of individual walkers. It is very obvious that the size of the walker is not used at all. This might be an artifact of our particular experiment. Our walkers were presented one at a time and the displays did not provide much reference to compare them to. If the absolute size of a walker could have been estimated more accurately from the displays, removing information about size would probably have resulted in an impaired identification performance. Both the shape of the walker as well as walking frequency seemed to provide diagnostic information. Depriving subjects from either of them resulted in small but significant
recognition drops in the learning curve. The results from the type I test sessions imply that both cues are rather independent and additive (Fig. 4.4B).

One of the main findings of this study is the fact that even after the deprivation of information about shape (including size) and walking frequency, recognition is still at a level which is about five to six times higher than chance level. The decreases in recognition performance due to normalization for shape and frequency are very small compared to the overall performance. Because deprivation of structural information has only a small effect, most of the information used for person identification seems to be extracted from the kinematics of the motion. The only parameter within this domain that was tested in the current study was gait frequency. Even though this parameter does affect performance, it does not seem to play a very major role. The second study therefore addresses the question of the importance of these particular aspects of human gait for person identification.
5. Study 2

5.1 Introduction

In the first study it was found that shape, but not size, affected the recognition of point-light walkers. The only kinematic cue that was normalized was the gait frequency. In the second study, the focus of attention was on the details of kinematic information by examining the relevance of single harmonics in the gait pattern (Experiment 1), as well as the impact of normalizing amplitude- and phase-spectra (Experiment 2).

The main question of Experiment 1 was whether all the relevant information about a walk is contained in the first harmonic, which explains the largest part of the overall variance in a walking pattern, or whether higher order harmonics are sufficient or necessary for person identification. If the amount of variance explained by the harmonics is crucial for identification, the performance should be high when the first harmonic is included in the model, but not above chance level when only the second or higher order harmonics are used. An alternative hypothesis is that the first harmonic represents primarily a general human walking pattern, and that individual information is contained in higher order harmonics. In this case, we should find a better identification performance in conditions when only the second or even only higher order harmonics are used in the displays of the walkers.

A further question concerns the necessary stimulus conditions to learn to discriminate individual point-light walkers. The stimuli used at the start of Study 1 were not normalized but contained their original information. Thus, it is not clear whether structural cues are necessary to learn to discriminate the walkers, or whether this task can be solved solely from the kinematics of the gait patterns. Therefore only point-light walkers that were normalized with respect to size, shape, and gait frequency were used throughout the complete experiment.

In both experiments observers were trained to name seven male walkers, whose individual gait patterns had been recorded beforehand. In the first experiment, the training part was followed by a non-reinforced test session in which the stimuli were normalized with respect to different harmonics, and a
final test session in which the viewpoints under which the walkers were shown were additionally varied. The second experiment followed the same design, with the exception that the test stimuli were normalized either for the amplitude or the for phase spectra of the walkers, and that only one final test session was conducted.

5.2 Experiment 1

5.2.1 Methods

Participants:
Eighteen subjects (9 female, 9 male, age 21-40 years, mean 25.89 years) participated in the experiment. All of them were students at the Ruhr-University who received course credit for the participation. Their visual acuity was normal or corrected-to-normal.

Stimuli and Materials:
When the second study was conducted, a larger database of walkers was available than for the first study. From this database of 40 walkers, 20 were chosen to serve as stimuli in the experiments. The ten most distinctive and ten least distinctive walkers from the database were excluded for this study. Distinctiveness was defined in terms of the Euclidian distance of each walker from the overall average. For each particular experimental subject, seven walkers were randomly selected from the set of the remaining 20 walkers. In this way it was possible to use the natural variance in the database without making the task too easy or too difficult.

All walkers were first normalized to unit size. The data were then subjected to a discrete Fourier analysis. Figure 5.1 shows the amplitude of the different harmonics averaged across all 20 walkers. Because the amplitude declines rapidly, the Fourier expansion was clipped after the fifth harmonic. The time invariant component of the Fourier expansion was then replaced by the average of this component computed across the seven walkers used for a
particular observer. This resulted in all point-light walkers having the same overall body structure. Finally, all walkers were normalized to the same average gait frequency. These normalized walkers (in the following called “BASE”-condition) were used in all training sessions. This is the same degree of normalization that was used for the final sessions in Study 1.

Modified versions for the test sessions were constructed in the following way: First the average across the seven walkers used for a particular observer was computed. For the first test set the first harmonic of this average walker was replaced by the respective first harmonic of each individual walker. Thus, seven different walkers were created which differed only with respect to the first harmonic (called “H1”-condition). Therefore, any possible differentiation between the walkers can only be caused by the perceptual significance of the individual first harmonic of each walker.

Comparable procedures were applied to create test sets of walkers which differed only with respect to the second harmonic (“H2”-condition) or to the third to fifth harmonics (“H3”-condition). The third to fifth harmonics were grouped, because their power was very small and visual analyses of the walkers suggested that there was virtually no difference between them if only one of these harmonics was used.

The displays were shown on a 19” screen (Iiyama Vision Master Pro 450), and the subjects were seated in front of the screen at a distance of 70cm.
The visual angle of the displays subtended about 4.8° in height and 2° in width under maximum excursion. Stimuli were rendered online at a rate of 85 Hz which was synchronized with the graphics board’s screen refresh rate.

**Design and Procedure:**

During training and the first test session, each observer saw the point-light walkers always from the same viewpoint. Equal numbers of observers were shown displays from a frontal view (FV, 0°), from a half-profile view (HV, 30°), or from a profile view (PV, 90°). Participants were randomly placed into one of three training groups according to these viewpoints, but the gender of the subjects was balanced within the groups. No occlusion was used in the displays, thus every marker was visible from all the viewpoints at all times. The experiment was conducted on two consecutive days for each participant. On the first day there were four training sessions, in which the observers learned to name the seven walkers. Only original walkers (BASE-condition) were used in the training sessions, so that the observers learned to name point-light walkers which were normalized with respect to size, shape, and gait frequency from the beginning of the experiment. In each trial one of the walkers was shown to the observer for 3 s. After the display, seven buttons with male names appeared in alphabetical order at the top of the screen, and the subject had to click on one of the names. If the answer was incorrect, the button with the correct name changed its colour for 1 s. If the answer was correct, nothing happened before proceeding to the next trial. In each session there were 140 trials, grouped into 10 blocks. Every walker was shown twice in each block, with the order of the walkers randomized within blocks. There was an inter-trial-interval of 1 s. After each complete training session there was a break of at least two minutes, in which the subject was allowed to leave the room.

Test Session A followed immediately after the four training sessions of the first day. In this session the BASE-walkers as well as the H1-, H2-, and H3-walkers were shown to the participants. The procedure was the same as in the training sessions, except that no feedback was given to the observers. Test Session A consisted of 140 trials (7 walkers x 4 types x 5 phase-randomized repetitions). The order of the trials was randomized across the whole session.
On the second day two more training sessions were conducted to refresh the participants’ knowledge of the walkers. After that, Test Session B was conducted. The procedure and the stimuli were the same as in Test Session A, but this time all three possible viewpoints were used. A total of 168 trials were conducted (7 walkers x 4 types x 3 viewpoints x 2 repetitions). The order of the trials was randomized across the whole session.

5.2.2 Results and Discussion

Analysis of the point-light walkers:
On average across the 20 walkers, the power, that is, the contribution to the overall variance of the first five harmonics of the walkers were 91.2%, 6.4%, 1.4%, 0.7%, and 0.3% of the overall variance, respectively. The square root of the power, that is, the amplitude of the Fourier components, scaled in terms of marker displacement averaged across the 20 walkers, time and the 15 markers is plotted in Figure 5.1. As can be seen from this figure, the average amplitude of the first harmonic is 414 mm, with a standard deviation of about 60 mm. Both values are about four times as high as the respective values for the second harmonic, and much higher than for the other harmonics. This indicates that a large part of the overall motion is expressed by the first harmonic, and that the differences between the walkers are also greatest for this harmonic.

No significant differences were found between male and female participants for any of the dependent variables neither during training nor in the test sessions, so the results presented here are pooled over both groups.

Training Sessions:
Starting from about chance level, all three groups quickly improved their performances in identifying the walkers (Fig. 5.2). In the last training session of the first day (i.e., in training-blocks 31-40), the FV-, HV-, and PV-groups reached an average correct response level of 76%, 71%, and 71%, respectively (mean: 73%). In the final training session (i.e., in training blocks 51-60, at the end of the second day), the performances reached an average level of 86%, 78%, and 87%, respectively. Paired t-tests showed that the differences between
the fourth and sixth training sessions reached significant levels in the FV- and the PV-groups \((p < .03)\), but not in the HV-group \((p = .103)\). The overall performances of the groups, averaged over all six training sessions, were 68%, 63%, and 65%, respectively. Univariate ANOVAs showed that neither the comparisons between the groups in the fourth or sixth training sessions nor the overall performances reached statistical significance (all \(p > .05\)).

Observers were quickly able to discriminate seven different walkers, which they were previously unfamiliar with, even though information about size, shape, and gait frequency were removed from the dataset. The walkers who were shown to each subject were selected on a random basis, so this finding cannot be caused by some walkers which were identified easily. Additionally, by choosing the walkers to have a medium distinctiveness, the possibility was ruled out that some subjects had to discriminate walkers who were very different from each other. The identification rate at the end of the training sessions was at about 80%, which is comparable to the findings of Study 1. This means that in the first study, high identification levels of normalized walkers were not caused by the fact that the walkers were shown in the original way beforehand, and that learning effects were transferred to subsequent sessions.
The performance of the HV-group saturated after four training sessions, whereas the levels of identification in the other two groups significantly increased over the last two sessions. Although the identification rates of the FV- and the PV-groups do not at any time differ significantly from the HV-group, there is a tendency for the latter group to perform worse than the others, especially in the last training session.

**Test Session A:**
A univariate repeated measures ANOVA with type of information (INFO, with levels BASE, H1, H2, and H3) as within-subjects factor and training viewpoint (VIEW, with levels FV, HV, and PV) as between-subjects factor was conducted, with the number of correct identifications of the walker-stimuli as the dependent variable.

The factor INFO had a significant effect on walker-identification \[F(3,45) = 64.73, p < .001\]. Performance deteriorated with the order of harmonic(s) that were used in the displays (Fig. 5.3A). Bonferroni-corrected Post-Hoc tests revealed significant differences between every single condition (all \(ps < .02\)). In the BASE-condition the displays were the same as those shown to the observers in the training sessions. The level of correct identifications in this condition (70%) did not differ from the average level of the previous training session (73%) \([t(17) = 1.272, \text{n.s.}\]). In the other three conditions the levels of identification (50%, 29%, and 20%, respectively) were still above chance level (paired \(t\)-tests, all \(ps < .03\)). However, if the value for the H3-condition was corrected for the number of conducted tests, it did not reach significance anymore.

There was neither a significant main effect for the factor VIEW \([F(2,15) = 1.07, \text{n.s.}\]) (Fig. 5.3B), nor for the interaction of factors INFO and VIEW \([F(2,15) = 1.41, \text{n.s.}\]). Figure 5.3C shows that the patterns of identification remained the same in all of the three groups. Although there was no overall interaction between the factors, a closer examination of Fig. 5.3C suggests that there could have been significant differences between the groups at single INFO levels, especially at the H1-condition for the HV-group. To control for these effects, additional Post-Hoc Tests were conducted between all groups for every single
condition, but none of these comparisons reached statistical significance (all ps > .05).

Figure 5.3: Identification performances in Test Session A: results for the type of information (A), Training Viewpoints (B), and interaction between the factors (C). Error bars indicate standard error of the mean; arrows indicate chance level.
The lack of differences between the groups in the training sessions as well as in Test Session A corroborates the results of Study 1, where it was found that the normalization of body structure affected identification performance from a frontal view most. No structural cues, which could have been used by the FV-group, remained in the displays of the walkers in the present study. Structural information, however, is not necessary for the identification of a person, because all groups performed well above chance level in the training sessions as well as in the BASE-, H1-, and H2-conditions of Test Session A. On the other hand, the performances obtained in the current study never reached the same levels as in Study 1, in which identification in the FV-group approached 100% if structural cues were available, whereas performance in the current experiment saturated at 80-90%.

Whereas the levels in the BASE-condition of Test Session A were comparable to the respective levels in the fourth training session, performance deteriorated as soon as the subjects encountered walker-displays deprived of parts of the kinematic information. However, performance in all groups was still above chance level in all conditions, although the results for the H3-condition do not hold out against the correction for the number of tests conducted. The results clearly indicate that the first and second harmonics provide sufficient information to adequately represent a walker in an identification task. Regarding the hypotheses, one can say that neither the absolute amount of variance explained by a harmonic nor its variability across walkers is necessarily predictive of the information used for person identification. The first harmonic explains on average about 91% of the variance in an individual’s walking pattern from the dataset used in this study, and it is the most important single factor of kinematic information that was varied, with an average of 50% correct identifications (i.e., more than three times higher than chance level) under the H1-condition. However, the performance in this condition is significantly lower than in the BASE-condition, showing that some important information is already missing. The second harmonic explains little more than 6% of the variance, but observers are still able to correctly discriminate the walkers at about 30% of all trials. The results of the H1- and the H2-conditions strongly support the hypothesis that the second harmonic also contains significant information for person identification.
One can also reject the hypothesis that the first harmonic describes a general human walking pattern, because otherwise observers would not have been able to discriminate the walkers by seeing only the first harmonic as individual information. The first harmonic comprises mainly the horizontal movements of the legs and the arms, as well as the lateral sway of the upper body, whereas the second harmonic contains a relatively big portion of the vertical movement in a gait pattern. The amplitudes of horizontal and lateral movements are much larger than the amplitudes of vertical motion. If the amount of motion determined the significance of a component for person identification, there should have been an advantage when the walkers had been shown from a profile view, which is the best viewpoint to perceive the most concise motions. Furthermore, one should have seen an interaction between the viewpoint and the information condition, with the performance under the frontal view being worse than from the other two viewpoints for the H1-condition, which was not the case.

Test Session B:
In this test session, one additional within-subjects factor was introduced, namely the test-viewpoint of the stimuli (TEST, with levels FV, HV, and PV). The remaining factors were the same as in Test Session A: VIEW (between-subjects, 3 levels) and INFO (within-subjects, 4 levels). A repeated-measures ANOVA was conducted with these three factors on the correct identification rate.

In accordance to Test Session A, a highly significant main effect was found for factor INFO $[F(3,45) = 50.80, p < .001]$, showing decreasing performance with higher order harmonics (see Fig. 5.4A). Multiple $t$-tests revealed that all conditions were significantly above chance level (all $p$s < .03). Again, when the $p$-values were corrected for the number of tests conducted, the value of the H3-condition did not reach a significant level anymore. Pairwise comparisons ($p < .01$) between the conditions revealed significant differences for all but the comparison between the H2- and the H3-condition. No effect was found for factor VIEW $[F(2,15) = 0.21, \text{ n.s.}]$ (Fig. 5.4B), but there was a significant main effect of factor TEST $[F(2,30) = 4.04, p = .028]$ (Fig. 5.4C).
Testing with the profile view produced a slightly worse performance than testing with the half-profile view.

![Graph A](image1.png)

![Graph B](image2.png)

![Graph C](image3.png)

Figure 5.4: Identification performances in Test Session B: results for the Type of information (A), Training viewpoints (B), and Test viewpoints (C). Error bars indicate standard error of the mean; arrows indicate chance level.
In the light of the results of Study 1, the advantage for the half-profile test view is rather surprising. The half-profile view is between the other two viewing angles, so one might suppose that there is a general advantage for this test viewpoint, because the displays still contain some of the information that was perceived from a more extreme view, and are thus more similar to the familiar view that was used in the training sessions. On the other hand, there was no advantage for the half-profile view over the frontal viewpoint. Maybe the effect is due to an asymmetry between the frontal and the profile view which seems to appear in Figure 5.5: generalization from the profile view to the frontal view seems to be easier than generalization from the frontal view to the profile view (compare the black and the white bars of the FV- and PV test viewpoints). Figure 5.5 also indicates a significant two-way interaction between the training- and the test-view, which was confirmed by the analyses \( F(4,30) = 17.36, p < .001 \). Performance in Test Session B was always best if the observers saw the displays from the same viewing angle as in the training sessions. No interaction effects were found for the other two combinations of the factors in Test Session B (INFO x VIEW and INFO x TEST).

![Figure 5.5: Interaction of factors VIEW (bars) and TEST (x-axis) in Test Session B. Error bars indicate standard error of the mean; the arrow indicates chance level.](image)

The interaction between the factors VIEW and TEST corroborates the results of Study 1. This effect is clearly caused by a benefit for test views which were familiar to the training groups, whereas the performance under the two views which were unfamiliar to the respective groups did not differ from each other.
This shows that the identification performance does not deteriorate linearly with an increasing amount of rotation between the training- and the test view. On the contrary, there is already an obvious forfeit after a rotation of 30° from the frontal to the half-profile view (45% vs. 31% recognition). It remains to be seen whether a rotation of less than 30° affects performance, or whether there is a discrete border of rotation which leads to a drop of performance when crossed.

There was also a significant three-way interaction between the factors in Test Session B \(F(12,90) = 3.60, p < .001\). Figure 5.6 shows that there is a similar pattern of identification performance over all INFO-conditions, but the gradient between the TEST-conditions becomes shallower with the higher order harmonics.

In this first experiment, the walking patterns were split up to examine the contribution of the discrete harmonics to the identification of the walkers. The first harmonic is most important for this task, but the second harmonic contains much more information than would be expected from its contribution to the overall variance of the walking pattern itself as well as from its contribution to the variability between walkers. Another way to represent the kinematic information of a walking pattern is to describe it as the amplitude- and the phase-values of the single markers. Whereas the amplitude spectrum determines the extent of displacement of the 15 markers, the phase spectrum contains information about the temporal relation between the movements of the markers and their spatial components. The differential roles of these two spectra were investigated in Experiment 2.
Figure 5.6: Interaction of factors VIEW (bars) and TEST (x-axis) under the four different INFO conditions (panels). Error bars indicate standard error of the mean; arrows indicate chance level.
5.3 Experiment 2

5.3.1 Methods

Participants:
Eighteen subjects (13 female, 5 male, age 19-42 years, mean 26.11 years) participated in the experiment. Because previous findings did not suggest any sex-differences in the performances of the observers, the number of male and female subjects was not balanced in this study. All participants were students at the Ruhr-University and were either awarded credits for completion of their psychology courses, or received 20 Euros for the participation.

Stimuli and Materials:
The same sample of 20 walkers, normalized by size, shape, and gait frequency, as described in the first experiment, as well as the same equipment was used. Seven walkers were randomly selected as target stimuli for each subject in the training sessions. For the final test session, the vectors encoding amplitude or phase for all harmonics of a single walker were replaced by the average values computed across the seven walkers used for the respective subject. Two sets of modified walkers were created: In the “AMP”-condition, walkers retained their individual amplitudes, but the vectors encoding phase were replaced with the averaged versions; in the “PHA”-condition, the walkers retained their individual phase-spectra, but the vectors encoding amplitude were replaced with the averaged versions.

While a custom designed C program was used for Experiment 1, the second experiment was conducted with Matlab and the Psychophysics Toolbox extension. Moreover, the size of the stimuli was somewhat different from the sizes used in Experiment 1. Walkers in the current experiment subtended a viewing angle of 8.1° in height and 3.3° in width under maximum excursion.

Design and Procedure:
The design of the second experiment was the same as that of the first experiment, with only few exceptions. The observers did not complete a test
session at the end of the first day. In the first experiment, Test Session A served
to collect more data about the role of the harmonics, independent of the test
viewpoint. Because comparable results were found in both test sessions of the
first experiment, the first test session in the second experiment was omitted.
The feedback procedure was slightly adjusted: when the subject clicked on the
correct name for a walker, the button flashed green for one second; when the
subject clicked on the wrong name, the button with the correct name flashed
red. At the end of the second day observers completed a single test session.
Stimuli were either the original walkers ("BASE"), as used in the training
sessions, or walkers from the AMP or PHA stimulus sets. All three viewpoints
(FV, HV, and PV) were used as possible viewing angles of the displays. The
test session consisted of 189 trials (7 walkers x 3 conditions x 3 viewpoints x 3
repetitions).

5.3.2 Results and Discussion

Training Sessions:
Comparable to the findings of Experiment 1, identification rates started from
about chance level, but quickly improved over time (Fig. 5.7). In the last training
session the FV-, HV-, and PV-groups reached mean identification rates of 89%,
92%, and 75%. The overall values over all training sessions were 73%, 77%,
and 59%. A comparison with the learning curves of the first experiment
suggests that whereas there were no obvious differences between the groups in
the first experiment, the FV- and the HV-groups showed higher performance
levels than the PV-group in the second experiment. However, t-tests failed to
reveal significant differences either for the overall performances of the groups or
for the performance in the last training session (all ps > .05). An examination of
the data of single subjects showed that the lower mean values in the PV-group
are due to one single participant who performed exceptionally bad.
A repeated-measures ANOVA on the rate of correct identifications was conducted with the factor Training Viewpoint (VIEW, levels FV, HV, and PV) as a between-subjects variable, and the factors Test Viewpoint (TEST, levels FV, HV, and PV) and Information (INFO, levels BASE, AMP, and PHA) as within-subjects variables. No effects were found for the factors VIEW \([F(2,15) = 1.49, \text{n.s.}]\) or TEST \([F(2,30) = 0.91, \text{n.s.}]\) (Figs. 5.8A and B). A highly significant main effect was found for the factor INFO \([F(2,30) = 96.45, p < .001}\) (Fig. 5.8C). Bonferroni-corrected Post-Hoc Tests revealed significant differences between all INFO-conditions, with the original stimuli leading to the best results (58.4%), followed by the stimuli that contained only information about amplitude (46.7%). Stimuli with phase-only information resulted in the worst identification (23.3%). All three values, however, differed significantly from chance level, as revealed by multiple \(t\)-tests (\(ps < .001\)).

The missing main effect for the test view is in contrast to the results of Experiment 1, where it was found that the half-profile test view resulted in a slightly better performance than the profile view. One can assume that the previous finding was a mere artefact, and it is concluded that there are no
Eminent differences between the test viewpoints, which is also in accordance with the findings of Study 1.

![Graphs showing main effects for the test session of Experiment 2: results for Training viewpoint (A), Test viewpoint (B), and Information condition (C). Error bars indicate standard error of the mean; the arrows indicate chance level.](image-url)

Figure 5.8: Main effects for the test session of Experiment 2: results for Training viewpoint (A), Test viewpoint (B), and Information condition (C). Error bars indicate standard error of the mean; the arrows indicate chance level.
With respect to the information condition, the results show that the contribution of the individual amplitudes of the markers of a point-light walker is far more important than the relations of the dots to each other, represented by their phases. To a certain degree, however, it is possible to identify a walker by the phase-spectrum of the gait pattern alone. Nevertheless, a level of 23% could already be reached if one of the seven walkers which were randomly selected for each subject had a very characteristic walk and could therefore be identified with a high accuracy. Therefore the walkers which were selected as the target stimuli for each subject were cross-checked, together with the respective identification rate under each condition, but no evidence was found that there were specific walkers which could be identified easily by all the observers for whom they had been used. It is therefore concluded that the phase spectrum of a walker has an effect on his identifiability, even though it is far less important than the amplitude. After subtraction of the chance level rate, the sum of the recognition rates obtained in the AMP- and PHA-conditions are about the same as in the BASE condition. Amplitude and phase spectra of a walking pattern therefore seem to contribute rather independently to the individuality of a walker.

As in Experiment 1, a highly significant interaction effect between the factors VIEW and TEST \(F(4,30) = 17.84, p < .001\) was found, which is shown in Fig. 5.9A. Again, performance was best if the walkers were shown from the same viewpoint as in the training sessions, and decreased with the amount of rotation away from the training viewpoint.

Another significant interaction was found between the factors TEST and INFO \(F(4,60) = 3.00, p = .025\). Additional repeated measures ANOVAs were conducted on the differences between the information conditions, separately for each test viewpoint. From a frontal view, all three information conditions differed significantly from each other \(ps < .005\). The best performance was found in the BASE-condition (58.8%), followed by the AMP-condition (40.8%) and the PHA-condition (25.7%). For the half-profile view, there was no difference between the BASE- and the AMP-conditions (60.1% vs. 50.9%), but performances under both conditions were significantly better than in the PHA-condition (23.8%; \(ps < .001\)). The same significance pattern was found for the
profile view, with 56.3% identification rate under the BASE-condition, 48.5% for the AMP-condition, and 20.6% for the PHA-condition (Fig. 5.9B).

The phase spectrum of a gait pattern only seems to be utilized from a frontal view, because performance dropped significantly when it was removed from the displays under this viewpoint (AMP-condition). In a point-light display, the dots which represent the arms and the legs of a walker convey the highest amplitudes. However, the full amount of the distances these dots cover cannot be seen from a frontal view. Therefore, the information about the amplitudes can only be partly used, and the phase spectrum of a gait pattern obtains greater importance relative to the other viewpoints. Probably the interaction between the factors is caused by this lack of information about the amplitude spectrum from a frontal viewpoint.

Figure 5.9: Interactions in the test session of Experiment 2 between factors VIEW and TEST (A), and between factors INFO and TEST (B). Error bars indicate standard error of the mean; arrows indicate chance level.
Furthermore, the analysis revealed a significant 3-way interaction between the factors VIEW, TEST, and INFO \([F(8,60) = 6.74, p < .001]\). The results for the BASE- and AMP-conditions resemble the findings which have already been shown in Fig. 5.9A. This interaction is levelled out when the amplitude of the walking patterns was averaged, because of the low overall performance under this condition (Fig. 5.10).

![Figure 5.10: Interaction of training viewpoint (x-axis) and test viewpoint (bars), separately for each information condition (panel). Error bars indicate standard error of the mean; arrows indicate chance level.](image-url)
5.4 Summary and Discussion of Study 2

The second study focused the attention on how person identification from biological motion is influenced by kinematic parameters, namely the harmonics as well as the amplitude- and the phase-spectra which result when a walking pattern is decomposed using Fourier analysis. This was done for three different viewpoints: frontal, half-profile, and profile view. Observers were able to learn to discriminate seven male walkers, independently of the viewing angle under which the training was conducted. It was found that the first and second harmonics both provided sufficient information to identify individual walkers. The task could also be solved above chance level when only the amplitude or only the phase information remained in the displays, although the former led to significantly higher performance levels. Removing information about the phase-spectrum impaired recognition much more when the walkers were shown from a frontal viewpoint, compared to the other two viewing angles. There was no structural information left in the displays that could be used for the task, meaning that kinematic parameters are sufficient to name the walkers.

Other studies (Mather & Murdoch, 1994; Troje, 2002a) found an overall advantage for the frontal view, which could not be replicated here. This was most probably caused by the lack of structural information in the current displays. Whereas the point-light displays were normalized with respect to size, shape, and gait frequency in Study 1, thus enabling the observers to first discriminate the walkers under a full information condition, displays which were normalized with respect to all three parameters were used throughout the complete second study. From the fact that the observers in both studies were able to solve the task at about the same accuracy levels, one can conclude that information about body structure and gait frequency is only of minor importance for person identification from biological motion. It was not explicitly tested how performance is affected when the complete kinematic information is normalized and only the structure of the walker remains unchanged. From the previous results one might assume that identification from structural cues alone leads to much lower accuracy levels. On the other hand, observers might adopt a
different strategy to discriminate gait patterns if only structural cues are available, so that differences in body structure are more readily perceived.

The first and second harmonics are both employed for the identification task: an elimination of one of these harmonics from the displays resulted in a significant drop of the recognition rate in relation to the original stimuli. In contrast, a combination of the third to fifth harmonics was not enough to obtain performances above chance level. Although the second harmonic explains on average only about 6% of the overall variance of a gait pattern from the dataset, it is nevertheless important for the recognition of individual walkers. Hence there is a clear difference between the physical walker-space that is dominated by the first harmonic, and the perceptual space, to which both harmonics contribute significantly. Higher order harmonics do not seem to provide much additional information. Even though this was not tested explicitly, one can assume that a second order Fourier representation is perfectly sufficient to represent an individual's walking pattern. Higher order contributions may simply encode measurement noise.

The amplitude-spectrum of a gait pattern is more important than the phase-spectrum, with the latter only being relevant primarily when the stimuli are shown from a frontal perspective. To a very limited degree, the relative phases of the markers are also used to identify point-light walkers from a half-profile and from a profile view. It might be instructive to examine whether there are specified sets of markers which lead to this identification from the phase-spectrum, for example, the forward swing of one arm in relation to the movement of the legs.

This is in close connection to the question whether it is necessary to see the whole body for a successful accomplishment of an identification task, or whether certain body parts might be sufficient. Kozlowski and Cutting (1977) suggested that even the ankle movements of a walker are sufficient to recognize the sex of the person, although this assumption was corrected later (Kozlowski & Cutting, 1978). Hill, Pollick and colleagues (Hill & Pollick, 2000; Pollick, Lestou, Ryu, & Sung-Bae, 2002; Pollick, Paterson, Bruderlin, & Sanford, 2001) showed that observers are able to estimate the affect of a person from arm movements, but that they are inefficient in estimating the gender of the
person from these kinds of motions. The role of specific body parts has not yet been examined in a person identification task.

A rotation of the displays between training and test sessions leads to a decline of performance. It is not clear which amount of rotation causes this deterioration, because there were no differences between views that were unfamiliar to the observers. This includes in particular the half-profile view, which is somewhere in between the other two views. One might assume that it is easier to generalize from or to this viewing angle, with respect to the other two, but this is clearly not the case. In the test sessions of the two experiments described in this paper only very coarse rotations of 30° or 90° were used, with respect to the frontal viewpoint. A finer gradation would be necessary to figure out critical levels of stimulus-rotation which impair the performance.
6. Study 3

6.1 Introduction

The first two studies investigated the relevant factors for learning and recognizing previously unknown gait patterns. In contrast, the aim of the third study was to examine the effects of visual and motor experience on a number of perceptual tasks. The independent manipulation of walking direction and playback direction allowed for the separate investigation of the effects of visual experience and motoric feasibility on the perception of walking patterns. Actors were recorded on a treadmill, walking forward or backward, and point-light displays of these events, which were played either forward or reversed, were shown to the very same observers who served as the actors. If visual experience enhanced perception of biological motion in general, point-light walkers apparently moving forward should be processed more efficiently than those moving backward. On the other hand, if perception is dominated by the motoric feasibility of a motion pattern, an enhanced recognition should be found for the displays which are shown in the normal playback direction.

Moreover, by presenting the displays to the very same subjects who served as the models for these displays, it was possible to contribute further results to the role of perception-action coupling. The common coding approach (Prinz, 1997) would suggest better recognition of self-generated actions than those of friends or strangers, because the perception of self-generated actions should lead to a higher similarity of activation patterns created by viewing and performing the same movements.

Two groups of female students participated in the experiment, and the members of the respective groups were known to each other. Therefore, it was possible to further examine the effect of visual familiarity with a gait pattern on its recognition: if the stimuli depicted the motions of members of their own groups, observers’ recognition performance should be better than when a motion pattern of the respective other group is used.

In four experiments, a variety of tasks that presumably act on different levels of biological motion perception were used throughout the experiments. In
the first experiment, the observers had to detect a walker that was camouflaged
by a scrambled mask. The second and third experiments investigated the
observers’ sensitivity for the naturalness of the displays. In the second
experiment, the task was to discriminate between point-light displays which
were played back either forward or reversed. In the third experiment only gait
patterns played back in the normal way were used, and the task was to adjust
the original step frequency of a walker under varying gait speeds. In the final
experiment the participants had to identify themselves and each other from their
point-light depictions.

6.2 General Methods

Participants:
Ten female first year psychology students (age 19-27, mean 21.9 years) from
the Ruhr-University served as actors for stimulus construction and as observers
for all experiments. They were recruited as two groups of five persons each,
who knew each other from their studies and spent at least ten hours per week
with the members of their respective groups. This recruitment served to create a
database of motion patterns of oneself, of a group of friends, or of a group of
“strangers”. The subjects received course credit for participation. All of them
had normal or corrected-to-normal vision. Table 6.1 indicates age, size, and
weight of the subjects.

| Table 6.1: Age, size, and weight of the participants of Study 3 |
|-----------------|----------------|----------------|
| Subject | Age (Years) | Size (cm) | Weight (kg) |
| 1 | 19 | 174 | 63.3 |
| 2 | 20 | 178 | 58.4 |
| 3 | 21 | 179 | 82.5 |
| 4 | 27 | 164 | 71.5 |
| 5 | 23 | 172 | 65.2 |
| 1 | 21 | 164 | 47.0 |
| 2 | 20 | 167 | 54.0 |
| 3 | 21 | 169 | 54.2 |
| 4 | 25 | 174 | 53.2 |
| 5 | 22 | 172 | 56.3 |
At the beginning of the recording session (see below), each subject orally answered a short questionnaire which was read to her. She had to estimate how much time per week she spent on average with the other members of her group. Furthermore, she indicated on a scale from 1 to 5 how well she knew each other participant of her group. The mean score on this scale in the first group was 3.05 (SD 0.78), and in the second group 3.55 (SD 0.48). The subjects reported to spend on average 20.2 (SD 7.7) and 19.3 (SD 6.2) hours per week with each other, respectively. T-tests revealed no significant differences between the groups in these two variables.

None of the participants reported having a medical condition, especially with the motoric system, that prohibited her from taking part in the experiment. The subject gave her written consent not to talk to other people, especially those of her own group, about the recording session or about the following perceptual experiments before the completion of the study. Because both groups of subjects consisted of first-year students, care was taken that both groups did not get to know of each other. If this had been the case, there would have been an increased probability that they had paid increased attention to the motions of the participants who served as the group of strangers in the perceptual tasks. On the other hand, it was a desired effect that the subjects paid more attention to the movements of the participants in their own groups, leading to a higher level of visual experience.

**Stimulus construction:**
The gait patterns were obtained with the subject walking forward or backward on a treadmill at different speeds. The respective speeds for forward walking were: 2, 3, 4, and 5 kph. For backward walking speeds of 2 and 3 kph were recorded. A backward speed of 3 kph was the maximum speed that could be applied without any participant having problems to stay on the treadmill. Because previous experience showed that actors exhibit a rather unnatural gait pattern when they are not familiar with a treadmill, the subject had to walk on it for 20 minutes before the start of the recording session, including forward and backward walking. Afterwards, the subject had to adjust the treadmill so that she walked at a comfortable speed. The mean self-adjusted walking speed was 4.17 kph (SD 0.56), averaged over all subjects.
In the recording session, the forward walks were captured first. The order of the speeds was randomized for all participants. For the backward walks, a speed of 2 kph was recorded first, to slowly familiarize the subjects with this condition again, followed by a speed of 3 kph. Data collection and the subsequent creation of the point-light stimuli followed the methods as described in Chapter 3.

In the following experiments, the point-lights were shown as white dots on a dark blue background. Three possible viewpoints were used: a frontal view, an oblique view, which was created by rotating the displays by 45° to the right, and a profile view with the walker facing to the right. The playback direction of the animations could be either normal or reversed, resulting in two apparent locomotion directions. An apparently forward moving walker could be created by an animation of a walker that was originally going forward on the treadmill, played in the normal way, or by a walker that was originally going backward, but played in reverse. These combinations result in a walker who is apparently moving in the same direction he is facing. On the other hand, apparently backward moving walkers were created by a forward going walker, played in reverse, or by a backward going walker played in the normal direction.

In the first experiment, the walkers were camouflaged by a mask of 120 dots, whose local motions were created of eight spatially and temporally scrambled versions of the same respective walker. Thus, the scrambled walkers always depicted the same walking direction and playback direction as the foreground walker, and they were shown from the same viewpoint. The offsets of these dots were randomly scattered across an area which covered 8.1° of visual angle in width and height at the centre of the screen. The position of the target walker inside this area randomly varied from trial to trial. If the display did not contain a walker, an additional set of 15 dots was added to the mask.

**Apparatus:**
The point-light stimuli were presented on an Iiyama Vision Master 19” screen, with a resolution of 1024 x 768 pixels. The subjects were seated in front of the screen at a viewing distance of 70 cm. The stimuli were shown using MATLAB and the Psychophysics Toolbox extension (Brainard, 1997). This apparatus and setup was used in all of the following psychophysical experiments.
6.3 Experiment 1

The goal of the first experiment was to examine whether visual or motor experience with walking patterns influence the detection of a walker in a scrambled mask of visual noise dots, and whether there is an enhanced sensitivity for one's own motion pattern in this kind of task. Previous studies which confirmed a higher sensitivity for self-generated than for other-generated actions (e.g., Knoblich & Prinz, 2001; Loula et al., 2004; Repp & Knoblich, 2004) only used tasks in which the agent of the action had to be named or discriminated. A generally enhanced sensitivity for displays of actions that were generated by oneself, however, should also show up in tasks which do not require explicit identification of the actor. A scrambled mask was used because it was shown to be of a sufficient complexity to reveal differences between individual conditions. The usage of a random dot mask would most probably lead to ceiling effects in all conditions.

6.3.1 Methods

*Design and Procedure:*  
Three to four weeks after the recording sessions, the participants conducted the first psychophysical experiment. They returned once every other week to complete the next experiments. Observers were seated individually in front of the computer monitor. They were told that there was another group of five persons who participated in the experiment and that they were only going to see some of those recordings shown as point-light displays. Because the task in the final experiment was to identify one’s own motions as well as those of the other members of the group, the reason for this was to keep the subjects from paying close attention to individual gait patterns and from establishing hypotheses about self- or friend-generated walks.

The target stimuli for each observer were chosen as follows: one stimulus was always the point-light display which depicted the observer herself (“self”). To select a stimulus from the group of the observer (“friend”), she was
asked whom she knew best of the other four group members, directly preceding the first experiment. She was told that this would be relevant for a different experiment in the course of the study. In the first group, two of the participants were reported twice as being best known to one of the other subjects. In the second group, every participant was reported once, thereby perfectly counterbalancing the frequency of self- and friend-generated gait patterns for each observer in this group. The last stimulus depicted a walker from the respective other group (“stranger”). Each participant’s point-light pattern served as a stranger-pattern for one of the observers of the other group.

In this experiment only the recordings from forward and backward walking at a speed of 3 kph were used. Point-light displays were generated which showed the walker from either a frontal viewpoint or from an oblique angle after rotating the walker by 45° to the right. Jokisch et al. (submitted) found an enhanced recognition of self-generated motion patterns from a profile view. However, showing the stimuli from the side would have resulted in a cluster of three dots each at the shoulder and at the hips, thus revealing the presence of a walker by its static configuration. Therefore, a rotation of only 45° was used. All walker stimuli were normalized to the same overall size. The height of the point-light walkers was approximately 4.1° of visual angle, and the width of the walkers under maximum excursion from an oblique angle was about 1.5°. They were camouflaged by a mask of 120 dots, which was constructed from eight scrambled versions of the same respective walker stimuli.

The animations were played back either forward or reverse. Of the overall 288 trials, half contained a target walker. In the 144 trials that did not contain a target walker (3 walkers x 2 gait directions x 2 playback directions x 2 viewpoints x 6 repetitions each), every combination of these conditions was used with equal frequency to create the mask. Every trial began with a small fixation cross that appeared for 500 ms in the centre of the screen. After this, a point-light display was shown for 2 seconds. The observers subsequently had to press a key in order to indicate if they had seen a human figure. This was followed by an inter-trial-interval (ITI) of 1 s. After every 72 trials an intermission of two minutes was inserted. The order of the trials was randomized. Sixteen training trials were conducted before the experiment proper started. The target
walker in these training trials was also chosen from the set of the other group, but not the same one that served as the stranger for the respective subject, so as not to familiarize the observer with any of the stimuli that appeared in the experiment. The completion of the experiment took about 30 minutes.

6.3.2 Results and Discussion

For each subject the percentage of correct responses was calculated. The overall detection rate was 78%, which was significantly above chance level of 50% \( t(9) = 8.31, \ p < .001 \). Additionally, the overall detection sensitivity was computed in terms of the \( d' \)-value (Macmillan & Creelman, 1991), which amounted to 1.58. The findings corroborate the results of previous studies (e.g., Bertenthal & Pinto, 1994; Pinto & Shiffrar, 1999) that local motion information is not necessary to detect camouflaged biological motion. Even with a set of 120 additional dots and a display time of 2 s, observers performed well above chance level. Whereas other studies used only profile views of a computer-generated walker, the present experiment showed that the task can be solved accurately from frontal and oblique viewing angles with a variety of individual walking patterns as well.

A 3 (walker type, ‘TYPE’) x 2 (gait direction, ‘GAIT’) x 2 (display direction, ‘DISP’) x 2 (viewing angle, ‘VIEW’) repeated measures ANOVA with detection rate as the dependent variable was used to analyze the data. No main effect for the factor TYPE was found \( F(2,18) = 0.36, \) n.s. \]. The observers were equally able to detect their own walking patterns (78%), their friends’ motions (77%) and strangers’ gait patterns (77%) in the scrambled mask.

Obviously, familiarity with a gait pattern, either motor or visual, had no effect on its detection in a scrambled point-light mask. The detection of a walker is an early task in the processing of biological motion. Previous studies (e.g., Knoblich & Prinz, 2001; Loula et al., 2004) had shown a familiarity effect on higher level tasks, including explicit identification of the agent or identity discrimination. The current results, however, can be interpreted as a first indication that this familiarity effect depends on the task at hand, and that the simple detection of a point-light walker is not sufficient to activate this effect.
Whereas there were no significant main effects for the factors VIEW \( F(1, 9) = 3.48, \text{n.s.} \) or DISP \( F(1, 9) = 1.90, \text{n.s.} \), a main effect for GAIT \( F(1, 9) = 8.43, p = .018 \) was found. When the walkers were going backward, they could be detected at a higher rate than when they were going forward (79% vs. 76%). This effect should be interpreted in the light of a DISP x GAIT interaction \( F(1, 9) = 6.95, p = .027 \). This interaction reflects the effect of the apparent motion, that is, whether the walker is moving in the direction he is facing. To ease the understanding of this effect, factors DISP and GAIT were collapsed into a new factor “apparent motion”. Figure 6.1 shows a slight advantage for the detection of walkers that were apparently moving forward (80% vs. 76%).

Moreover, there was a significant interaction between the factors TYPE, DISP, and GAIT \( F(2, 18) = 4.91, p = .02 \). Again, the factors DISP and GAIT were collapsed into the factor “apparent motion”, to clarify this result. Figure 6.2 shows that the detection of an apparently forward moving walker was only better for self-generated gait patterns (83% vs. 73%). There were no obvious effects of the apparent motion direction on the perception of friends’ or strangers’ gait patterns.
The main effect of the apparent motion direction indicated a role of visual familiarity with a gait pattern for its detection in a scrambled mask. Human observers are experienced in watching other people moving forward, whereas backward motions are hardly seen in our everyday environment. Thus, this visual experience could have explained the slightly higher saliency of apparently forward moving walkers. However, this effect was based exclusively on a difference for self-generated motions, contradicting the effect of visual familiarity. Instead, a combination of visual and motor familiarity would have to be assumed to explain this effect.

The first experiment did not bring conclusive evidence for an enhanced recognition of self-generated gait patterns over other-generated motions in a detection task. This is probably caused by the task at hand, because the detection of a coherent figure in a scrambled mask does not necessarily lead to the activation of the representations of a specific walker.

Moreover, there was no effect of the display direction, showing that a change of the trajectory forms of the point-lights did not impair the detection of the human structure in the mask. However, little is known about observers’ sensitivity for reversed animations of point-light events. Wickelgren and Bingham (2001) found that infants are sensitive for the playback direction of an event. On the other hand, the results of Pavlova et al. (2002) suggest that the visual system lacks sensitivity for reversed transformations of biological motion.
because observers often misinterpreted the playback direction of a reversed event. However, it was not directly examined whether the participants could discriminate animations which were played forward from those played backward. This was investigated in the second experiment.

6.4 Experiment 2

The same participants as in the first experiment were tested to examine their sensitivity for the naturalness of the displays. Again, they watched point-light walkers, created either of their own recordings, a friend’s gait pattern, or that of a stranger, walking either forward or backward, and played back in the normal way or reverse. Three different viewing angles were used. The observers’ task was to determine the playback direction of the display. This gave an estimate of their knowledge about how a natural walking pattern looks like, and whether this is affected by the viewpoint.

No differences between self- and other-generated actions were found in the detection task of Experiment 1, which was probably caused by the task itself, because the perception of a point-light walker in a scrambled mask is a rather early step in the processing of biological motion. It seems reasonable that the discrimination of natural and unnatural displays activates later stages of this process, which could in turn lead to a better discrimination of the playback direction for self-generated walking patterns than for those of friends or strangers.

6.4.1 Methods

Stimuli:
Each participant’s recordings of 3 kph, walking forward as well as backward, were used to create the point-light stimuli. The displays were shown either from a frontal view, from an oblique angle, or from a profile view, without visual noise. The point-light walkers for this experiment were not normalized, so that each
stimulus retained its original size. The overall size of the stimuli was on average 6.1° of visual angle in height and 2.9° in width under maximum excursion from a profile view. For each observer, the same walkers as in Experiment 1 were used to create the point-light stimuli of a friend’s or a stranger’s gait pattern. The animations of the point-light walkers were played either in the normal way or in the reversed direction.

**Design and Procedure:**
The observers were tested individually, and they were told again that they would see the point-light depictions of the respective other group. The experiment consisted of 288 trials (3 walker types x 3 viewing angles x 2 gait directions x 2 display directions x 8 repetitions each). An intermission of two minutes was inserted after every 72 trials. The course of each trial corresponded to the setup of Experiment 1. The task of the observers was to specify the display direction of each trial, that is, to indicate whether the animation was played back in the normal way or whether it was reversed. The participants completed twelve training trials before the beginning of the experiment proper. For each subject, the same walker that was used for the training trials in the first experiment was chosen again for this task. The completion of the experiment took about 30 minutes.

**6.4.2 Results and Discussion**

For each subject the percentage of correct identifications of the display direction was calculated. The overall performance of correct direction recognition was 60%, which was significantly above chance level of 50% \( t(9) = 2.96, p = .016 \). This performance corresponds to a \( d' \)-value of 0.56. Furthermore, an overall answer-bias was computed (Macmillan & Creelman, 1991), which resulted in a \( c \)-value of -0.34, indicating a slight tendency of the observers to answer “normal display direction”.

A 3 (walker type, ‘TYPE’) x 3 (viewing angle, ‘VIEW’) x 2 (gait direction, ‘GAIT’) x 2 (display direction, ‘DISP’) repeated measures ANOVA was used to analyze the data. The ANOVA revealed a significant effect for the factor
VIEW \( F(2,18) = 4.46, p = .027 \). Bonferroni-corrected Post-Hoc tests showed that performance from a profile view (62%) and from an oblique angle (63%) was better than from a frontal viewpoint (55%; both \( ps < .01 \)). Moreover, a significant effect of the factor DISP was found \( [F(1,9) = 45.80, p < .001] \). Stimuli that were played in the normal direction were discriminated with a greater accuracy than animations that were played in reverse (72% vs. 48%). Although this effect is at least partly based on the answer-bias of the observers, it seems reasonable to conclude that the observers were only confident in their estimations when the animations were played forward. Judgements about reversed animations, on the other hand, were at chance level.

Moreover, there was a significant interaction effect of the factors GAIT and DISP \( [F(1,9) = 24.58, p = .001] \). Comparable to Experiment 1, this reflected an effect of the apparent motion direction (Fig. 6.3). Discrimination rate was higher when the walker was apparently moving forward (68%) than when he was apparently moving backward (52%).

Again, this finding most probably reflects an effect of the visual familiarity with a gait pattern. When the walker was apparently moving forward, the observers could tell whether the gait depicted a natural locomotion pattern, because humans are experts in watching persons going forward. On the other hand, when the walker was apparently moving backward, the naturalness of the gait pattern could not be discriminated, due to the lack of experience.
Furthermore, a significant interaction between factors VIEW and DISP was found \( F(2,18) = 4.67, p = .005 \) (Fig. 6.4). Additional ANOVAs on the factor VIEW, separately for both display directions, showed that when the animations were shown forward, discrimination performance from an oblique angle was better than from the other two viewpoints \((ps < .001)\). In contrast, when the animations were shown in reverse, performance from a profile view was better than from the other two viewing angles \((ps < .001)\).

![Figure 6.4: Discrimination rate of forward and reversed animations, separately for each viewpoint. Error bars indicate standard error of the mean.](image)

After the completion of the experiment, each observer was asked whether she had a specific strategy to discriminate normal and reversed displays. A common answer was that the centre of mass was shifted aback in persons who were walking backward. This would be in accordance with the significant effect of the viewpoint, because this shift can only be detected from a viewing angle that noticeably deviates from the frontal view. To test for a possible shift of the centre of mass, the horizontal positions of the markers which represent the pelvis and the sternum in the sagittal plane were computed, separately for forward- and backward-going walkers. However, a paired \( t \)-test revealed no significant difference in the horizontal distance of these markers \([t(9) = 0.79, \text{n.s.}]\), indicating that the centre of mass did not shift with walking direction.

Other cues reported by the subjects were the amount of arm swing and the step length, with higher amplitudes indicating forward walking. Therefore, the amplitudes of the wrist- and ankle-markers in the horizontal direction were
computed for each walker, averaged over the respective left and right joints, separately for forward and backward going walkers. Paired $t$-tests revealed significant differences between the walking directions for the amplitude of the ankle-markers [$t(9) = 6.87, p < 0.001$], with the forward going walker showing higher amplitudes, but not for the markers at the wrists [$t(9) = 1.46, \text{n.s.}$]. Thus, the step length is apparently a reliable indicator of the walking direction, with persons walking forward taking longer steps. Figure 6.5 presents the average trajectories of the right ankle marker, moving either forward (black line) or backward (grey line). Whereas the average forward step covers about 50-55 cm, which is in accordance with values reported in the literature (e.g., Bertram & Ruina, 2001), a backward step is about 10 cm shorter. In addition, it is also shallower than a normal step.

![Figure 6.5: Trajectories of the right ankle marker of the average walker on a treadmill. The black line depicts walking forward, the grey line walking backward.](image)

Obviously, because the actors were walking at a constant speed (3 kph), the step length is in reverse relationship with step frequency. Therefore, step frequency would be another factor to discriminate walking direction. In the present data, the mean frequency of backward motion was 0.91 Hz, whereas the mean frequency for forward walking was 0.78 Hz. On the other hand, none of the observers reported that she had used the gait frequency to discriminate naturalness.
To my knowledge, the current study was the first to investigate the ability to discriminate the playback direction of biological motion. Although the results show that the subjects were capable of judging the naturalness of the displays, the overall performance of 60% indicates that it was a highly challenging problem, and it is not clear which cues were used by the observers to solve it.

Obviously, the frontal viewpoint does not offer a sufficient set of clues. In contrast, rotated displays of point-light walkers comprise enhanced information about the motion of the joints, mainly in the horizontal direction. The results indicate that the observers were able to pick up at least some of this kinematic information to solve the task at hand. Apparently, the overall amount of motion, as projected onto the fronto-parallel plane, is not of critical importance. Otherwise, performance under the oblique viewing angle should have been lower than under the profile view, because the main portion of the overall motion is oriented in the horizontal direction and can be wholly perceived only from a profile viewpoint. The lack of difference between these two viewpoints suggests a non-linear relationship between degree of rotation and performance.

No conclusive evidence was found that motor experience exerts any influence on the discrimination of the display direction. Moreover, familiarity with specific gait patterns, that is, those of friends, does not seem to play an important role. The effect of the apparent motion direction indicated an influence of the visual familiarity with gait patterns *per se*, because observers are experienced in watching persons who are going forward. Therefore, the third experiment used a different paradigm to further examine this sensitivity, namely the active adjustment of the natural gait frequency of a walker.

### 6.5 Experiment 3

Whereas passive perceptual tasks were used in the first two experiments of this study, requiring the observers only to answer two-alternative forced-choice questions, the third experiment needed a more active participation of the subjects. This time their task was to adjust the gait frequencies of the walkers until these were deemed to walk at their original frequencies. Instead of only
discriminating between two possible answers, this task intended to oblige the observers to deal more thoroughly with natural representations of biological motion.

6.5.1 Methods

Stimuli:
The recordings of each participant walking backward at 2 and 3 kph, and walking forward at 2, 3, 4, and 5 kph were used to create the point-light stimuli. The displays were shown either from a frontal view or from a profile view. The oblique viewing angle was not used because no significant differences between the oblique and profile view were found in the second experiment. The overall size of the stimuli was on average 8.1° of visual angle in height and 3.9° in width under maximum excursion from a profile view. For each observer, the same walkers as in the first two experiments were used to serve as the point-light stimuli of a friend’s or a stranger’s gait pattern.

Design and Procedure:
In each trial, one point-light walker was shown in the centre of the screen. The initial gait frequency of the walker was randomly determined from a range between 0.5 and 1.5 times the original frequency. For example, if the original frequency of a walker was 0.8 Hz, the initial frequency of the displayed walker was randomly determined between values of 0.4 and 1.2 Hz. On the left side of the screen there were eight buttons which could be used by the subject to adjust the gait frequency of the walker in several gradations. The top four buttons, labelled ‘Faster (+++)’ to ‘Faster (+)’, increased the frequency in steps of 0.167, 0.083, 0.042, or 0.008 Hz, respectively. The lower four buttons, labelled ‘Slower (-)’ to ‘Slower (----)’, decreased the frequency by the same amounts. The task of the subject was to adjust the frequency by repeatedly clicking the buttons until she thought that it was the natural frequency of the original recording, and then to click on a button labelled ‘OK’, in the top right corner of the screen. The next trial started after an ITI of 1 s.
The observers were told again that they would see the point-light depictions of the respective other group. The experiment consisted of 72 trials (3 walker types x 2 viewing angles x 6 gait speeds [4 forward, 2 backward] x 2 repetitions each). An intermission of two minutes was inserted after 36 trials. The observers completed six training trials before the beginning of the experiment proper, again with the data of the walker that was used for the training sessions of the previous experiments. The completion of the experiment took about 40 minutes.

6.5.2 Results and Discussion

For each trial the absolute difference between the original gait frequency and the estimated frequency of the walker was computed. Two repeated measures ANOVAs were computed to analyze the data, separately for both gait directions. For walkers going forward, a 3 (walker type, ‘TYPE’) x 4 (speed, ‘SPED’) x 2 (viewing angle, ‘VIEW’) ANOVA was conducted. No significant main effects or interaction effects were found. The standard deviation between original and estimated gait frequency was 0.21 Hz. For backward walking, another 3 (TYPE) x 2 (SPED) x 2 (VIEW) ANOVA was conducted. Again, no significant main effects or interaction effects between the factors were revealed. The standard deviation between original and estimated gait frequency was 0.26 Hz. To examine whether there was a difference in the quality of estimations between forward and backward walking, only speeds of 2 and 3 kph were analyzed, because no other speeds were recorded for backward walking. An additional repeated measures ANOVA showed that gait direction did not have a significant effect on the accuracy of the estimation \( F(1,9) = 1.09, \text{n.s.} \).

The observers were equally well able to adjust the original gait frequency of a walker under a variety of walking speeds, gait directions, and viewpoints. The missing effect of the gait direction indicates that visual experience does not play an important role for this quality. Similarly, if the familiarity with a gait pattern affected performance, there should have been an effect for factor SPED, with better estimations for gait speeds of 4 kph, because this was closest to the average self-adjusted speed of 4.17 kph on the treadmill. The
overall results might have been affected by the repeated usage of point-light walkers at a speed of 3 kph. In the first two experiments only walkers of this speed were used as stimuli, and the observers have possibly transferred the knowledge of these displays to the current experiment. This is in accordance with the finding that the most accurate estimations of gait frequency occurred for walkers of 3 kph (standard deviation of 0.2 Hz), although there was no significant overall effect of gait speed. Nevertheless, it would be instructive to obtain data from a naïve group of observers who did not have any experience with viewing point-light walkers before.

The viewpoint effect that was found in Experiment 2 did not show up this time. Apparently, the length of the steps, which was found to have an effect on discriminating display direction, and which could only be assessed from a profile view in the current experiment, is not a necessary factor in determining the gait frequency of a walker.

Another possible parameter for estimating the frequency is the overall motion of the body, with higher amplitudes indicating a higher speed and thus a higher step frequency. One might also argue that the observers used the size of the walkers to infer the frequency, because the displays were not normalized with respect to body size in the current experiment. However, in contrast to the study of Jokisch and Troje (2003) there were no pictorial cues indicating the absolute size of the walkers, so that only the relative sizes of the respective three walkers shown to each observer could have been estimated. Moreover, the participants were sensitive to a variety of frequencies displayed by the same walkers over several gait speeds, whereas size remained constant under these conditions. The same argument holds true for the static body structures of the walkers. Therefore, the observers most probably used kinematic factors which varied between gait speeds to estimate the original step frequencies of the walkers.

A standard deviation of 0.2 – 0.25 Hz from the original gait frequency appears to be an accurate estimation, thus indicating a high sensitivity of the observers to natural walking patterns. Unfortunately, this result might be confounded with the range of the initial frequency at the start of each trial. The initial frequency was not recorded, due to a technical mistake in the experiment. A comparison of the differences between the original and the estimated
frequencies, as well as between the original and the initial frequencies would have to be conducted to support the assumption that the results reflect a veridical sensitivity of the observers for the naturalness of the stimuli. A novel experiment with naïve observers is currently in preparation to answer this question.

6.6 Experiment 4

In the final experiment the task of the observers was to identify themselves and the members of their respective groups, that is, they had to explicitly name the walker that was shown on the screen. Previous studies that examined groups of persons known to each other found contradicting results about an enhanced sensitivity for self-identification from point-light displays. Cutting and Kozlowski (1977) did not find significant differences between the identification of oneself and that of friends. On the other hand, Beardsworth and Buckner (1981) reported a better recognition of self-generated locomotion patterns when these movements were slightly more complex and when the number of answers in favour of oneself and of the others was restricted according to the frequency of the occurrence of the walkers. Jokisch et al. (submitted) found an enhanced self-identification of point-light walkers seen from a profile view, but not from a frontal view or from an oblique viewing angle. Loula et al. (2005) reported a noticeable discrimination performance only for complex motion patterns, but not for locomotive movements like walking or running. In the last experiment of this study, the same participants as in the first three experiments had to identify their own gait patterns as well as those of the members of their respective groups shown as point-light displays. In addition to contributing further data to the role of the viewpoint in person recognition, the purpose of this last experiment was to investigate whether person identification is affected by the walking direction of the actor as well as by the playback direction of the display.
6.6.1 Methods

Stimuli:
The gait patterns of each participant walking forward and backward at 3 kph were used to create the stimuli. Each observer saw the point-light displays of the five members of her respective group either from a frontal or from a profile view. The displays were normalized with respect to overall size, and they were played back either forward or reverse. The average size of the stimuli corresponded to the values given in Experiment 3.

Design and Procedure:
The experiment consisted of 240 trials (5 walkers x 2 viewpoints x 2 walking directions x 2 display directions x 6 repetitions each), split up into six blocks of 40 trials. In every block, each combination of the four factors was used once in a random order. On every trial, one walker was shown for 2 s. Afterwards, the names of the group members appeared on the left side of the screen, and the participant had to click on one of the names in order to indicate which walker had just been shown. The next trial started after an ITI of 1 s. An intermission of 2 min was inserted after every 60 trials. The observers completed four trials with walkers from the respective other group and fictitious names to accustom to the procedure. The completion of the experiment took about 30 min.

6.6.2 Results and Discussion

A 2 (type of walker, ‘TYPE’) x 2 (viewing angle, ‘VIEW’) x 2 (gait direction, ‘GAIT’) x 2 (display direction, ‘DISP’) repeated measures ANOVA was conducted on the number of correct walker identifications for each observer. No significant main effects or interaction effects between any of these factors were found. Observers could identify their own gait patterns at a level of 30%, and those of their respective group members at a level of 28%. Both values are significantly above chance level of 20% (ps < .001). Note that chance level for self-identification as well as for friend-identification is both at 20%, because the observers’ task was to identify the individual persons, not to discriminate
between self and friends. Table 6.2 depicts the confusion matrix of the walkers, that is, the percentage of answers in favour of a walker depending on the walker that was shown.

Table 6.2: Percentage of answers for the respective walkers of each group in Experiment 4. Correct identifications are printed bold. Percentages in rows of the groups sum to 100%, deviations are due to rounding.

<table>
<thead>
<tr>
<th>Answer</th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5</td>
<td>1 2 3 4 5</td>
</tr>
<tr>
<td>1</td>
<td>22 26 18 15 18</td>
<td>40 12 15 15 19</td>
</tr>
<tr>
<td>2</td>
<td>38 33 10 8 12</td>
<td>21 37 25 8 10</td>
</tr>
<tr>
<td>3</td>
<td>10 20 19 18 33</td>
<td>13 22 34 15 16</td>
</tr>
<tr>
<td>4</td>
<td>15 15 21 25 24</td>
<td>11 13 23 25 28</td>
</tr>
<tr>
<td>5</td>
<td>11 23 31 16 18</td>
<td>16 9 14 33 28</td>
</tr>
</tbody>
</table>

Apparently, the performance in the second group (33%) was better than in the first group (24%). A \( t \)-test confirmed that there was a marginally significant difference in the number of correct identifications between the groups \( t(8) = 2.31, p = .05 \). It is unlikely that this effect was caused by a different degree of familiarity of the participants in both groups. The two parameters indicating the degree of familiarity in both groups (see section “Participants” in the General Methods of this study) had not revealed significant differences. However, it is possible that some participants got to know each other better during the 6-7 weeks in between the recording session and the completion of the last experiment, but it is implausible why this should have happened predominantly in the second group. Looking at the matrix in Table 6.2, one can see that there are two pairs of walkers in the first group that were often confused with each other, namely walkers 1 / 2 and 3 / 5. In both cases these walkers were mistaken more often for the respective other person than they were identified correctly. In the second group, there is only one such pair, namely walkers 4 and 5. Thus, the overall difference between the groups might have been caused by the higher difficulty of discriminating these two pairs of walkers in the first group.
Given the number of only five participants per group, a level of correct identifications of about 30% is relatively low, and the lack of significant effects indicates that this was not caused by the conditions of the experiment, that is, by the viewing angle, gait direction, or display direction. Cutting and Kozlowski (1977) as well as Beardsworth and Buckner (1981) reported overall levels of more than 35% correct responses with six walkers each. There are some differences between these studies and the current experiment that might have caused these discrepancies. First, the walkers in the present experiment were normalized with respect to their overall sizes, taking away some of the individual information that might be used for person identification in a natural environment. On the other hand, Study 1 of the current work had shown that body size does not seem to be a necessary criterion to solve an identification task from point-light displays. Moreover, the subjects of Cutting and Kozlowski (1977) were reported to be of approximately the same height and weight, whereas no specification about body shape can be found in the study of Beardsworth and Buckner (1981). Furthermore, in the study of Jokisch et al. (submitted), in which the data of the walkers were also normalized with respect to size, the levels of correct identifications varied between 18% and 30%, but with twelve walkers. Taken together, these findings can be interpreted as evidence that normalizing the size of a walker did not have a major effect on her identification in the current experiment.

A more important difference between previous studies and the current experiment is that the subjects were recorded walking on a treadmill. Previous experience from other recording sessions in our laboratory had shown that walkers on a treadmill exhibit a different gait pattern, compared to normal, unconstrained walking. The treadmill was used to keep the speed of all walkers constant, so that a specific gait speed did not indicate the identity of a walker. An identical speed of all walkers could have also been produced by adjusting each walker’s step frequency, because the walking speed is a function of step length and frequency. However, this would have changed the natural dynamics of the display. Moreover, the third experiment, as well as the results of Jokisch and Troje (2003), indicated that observers are sensitive to these natural dynamics, so that changing the dynamics of a walker might have interfered with
her identification. Therefore, only walkers that were as close as possible to the distal stimulus were used.

To enhance the familiarity of the observers with the apparatus, they had to practice walking on the treadmill for at least 20 minutes before the start of the recording session. Nevertheless, the recorded gaits probably deviated from natural walking patterns as obtained in everyday contact with the subjects. On the other hand, a profound deviation from a natural walking style should have become evident in a significantly impaired identification of the respective group members, compared to self-recognition. The identification of one’s own gait pattern would presumably be less impaired, because the participants could have memorized how their gaits were affected by the treadmill and recognized this behaviour when they saw point-light versions of their recordings.

6.7 Summary and Discussion of Study 3

The main purpose of the third study was to examine the effects of visual and motor experience with walking patterns on a variety of perceptual tasks. A general effect of motor experience should have manifested itself in an enhanced sensitivity for one’s own locomotion pattern. Similarly, a significant influence of visual familiarity would have been shown by a better recognition of a friend’s gait pattern as well as of those point-light stimuli which showed a figure walking forward, compared to a backward directed walker. Neither of these two effects could be demonstrated in any of the four experiments.

The participants were recruited as two groups of observers who were known to each other within their groups. Thus, the effect of visual familiarity was examined by comparing the sensitivity for walking patterns of persons of the observers’ own groups with that for the motions of the respective other group. Because all participants represented a sample of the same group of first-year students, care was taken that they did not get to know the members of the other group, to prevent them from paying close attention to the gaits of those subjects. Otherwise, one could not have assumed an increased visual familiarity
with the movements of the members of one’s own group over the familiarity with the members of the other group.

In the first experiment, the observers’ task was to detect a human figure within a scrambled mask. Whereas the average performance was clearly above chance level, no conclusive effects were found for any of the main conditions used in the task. Most importantly, there was no enhanced detection performance for walkers which represented the subjects’ own gait patterns or those of their friends, compared with the performance for detecting walking patterns created by members of the strangers’ group. This is a first indication that an enhanced sensitivity for specific motion patterns depends on the task to be performed. On the other hand, there was an effect of the apparent motion direction, indicating a role of visual experience. Nevertheless, a human structure can be perceived even if it performs an unfamiliar locomotion pattern (walking backward) or if it exhibits a movement that is biomechanically impossible (reversed display). The detection of a coherent figure under conditions of visual noise is extremely robust against a number of perturbations, as had been shown by previous studies (e.g., Ahlström et al. 1997). The accuracy of the detection of unfamiliar motion patterns had already been demonstrated by Shipley (2003), who showed that observers could recognize a person walking on his hands.

Moreover, the viewpoint from which this motion is shown did not affect its recognition in the first experiment. Previous studies only used profile views of the walkers in detection tasks. Apparently, this was mainly caused by the limited methods to create point-light walkers. Even contemporary studies still use the algorithm advanced by Cutting in 1977, which only covers a profile view. The present experiment was the first to indicate that the detection of a target walker in a scrambled mask is also possible from viewpoints that deviate from the profile view. Obviously, the projected amplitude of the motion, which increases with the amount of rotation against the frontal view, does not play a significant role for the task, because no significant differences were found between the performances under both viewpoints. On the other hand, the effect of the amplitude is counteracted by the local motion of the dots of the visual mask, because they resemble the individual trajectories of the dots which represent the joints of the walker. An effect of the viewpoint would probably show up if a
random dot mask was used. However, an indicator for an enhanced difficulty of the task is the overall level of performance. A $d'$-value of 1.58 is slightly lower than the levels reported in previous studies (e.g., Bertenthal & Pinto, 1994; Pinto & Shiffrar, 1999), where the $d'$-levels varied between values of 1.84 and 2.56. On the other hand, the proportion of mask dots to walker dots in the current experiment was higher than in those two studies, which is an alternative explanation for the performance differences. Obviously, the point-light walkers that are created by the method used throughout this work are highly salient, independent of the viewing angle. However, a direct comparison between this method and the algorithm advanced by Cutting (1977) is still needed.

Whereas there was no effect of the viewpoint in the first experiment, the second experiment revealed a significantly lower performance from a frontal view, compared to an oblique angle and a profile view, when the observers had to discriminate between animations played forward or reversed. It was found that the step length was apparently a reliable indicator for the true walking direction, which is obviously confounded with the gait frequency at a constant speed. Again, there was no evidence of an effect of motor experience on the task, because the type of walker did not have a significant influence on the discrimination performance. However, a significant influence of the apparent motion direction of the walker was found, supporting the general effect of the visual familiarity with gait patterns, as found in Experiment 1.

The overall performance for the task was rather low, corroborating findings of Pavlova et al. (2003) who demonstrated that reversed displays are often mistaken for different stimuli moving forward. Why is there such an obvious contrast between the ability to detect biological motion under visual noise, and the capacity to judge its original direction? Humans are experts in the perception of biological motion, but this expertise should bear on both, the detection of a human figure as well as on the discrimination of the playback direction. However, from an evolutionary point of view, the detection of an animate being played an important role for the individual's survival. On the other hand, watching reversed displays is an event that is never encountered in a natural environment, but that is only generated artificially. Therefore, the discrimination of the natural direction of a gait pattern is also an artificial task,
which does not occur in ecological settings. Thus, the performance differences might be based on the observers’ experience with the task at hand.

The estimations of the original gait frequency in Experiment 3 were rather accurate, although further experiments are needed to clarify this ability. Previous studies demonstrated the ability of the observers to derive dynamic parameters from kinematic patterns in the displays which depended on extraneous factors, such as a weight that had to be lifted. The current experiment, however, showed that the observers are sensitive to the dynamic factors of the gait pattern itself. A prominent factor to determine the frequency might be the overall motion of the gait pattern, especially the amount of arm swing, with higher amplitudes occurring at a high walking speed. The impact of the amplitudes could be examined by a procedure similar to the one that was used in the second experiment of Study 2. For example, if the amplitudes of the gait patterns play a significant role for the recognition of the original gait speed, performance should drop if these are normalized across all walking speeds.

In the final experiment the participants had to identify their own gait patterns and those of their friends in a multiple choice task. The overall level of correct identifications, compared to previous studies on this subject, was rather low under all conditions, although the average performance was still above chance level. This was probably caused by the subjects having to walk on a treadmill, instead of being allowed to walk freely across the room. No effects of the type of walker, viewing angle, gait direction or display direction were found. Apart from the study of Jokisch et al. (submitted), the current experiment was the only one that examined individual recognition of biological motion patterns from non-profile views. Whereas those authors found a better self-recognition from a profile view, compared to other-recognition, the current study did not confirm this result. The low overall performance in this experiment might have contributed to the lack of difference between self- and other-recognition. On the other hand, contradicting results can be found in the literature, some of them pointing to an enhanced self-recognition from biological motion (Beardsworth & Buckner, 1981), and some of them objecting this result (Cutting & Kozlowski, 1977).

The recognition of familiar people from point-light displays depends on numerous factors, which have to be controlled carefully. The familiarity with the
individuals is an important point – persons who are in everyday contact with each other are probably more sensitive for little details of a walking pattern than other persons, and thus more likely to recognize the person from such degraded displays. The participants of the current experiments were all recruited from the same sample of first-year students who were in contact with each other about five days a week. Although they had known the members of their groups for only about four months before the accomplishment of the recognition experiment, they were assumed to have at least a medium level of familiarity with the motion patterns of their respective group members. Moreover, the sole participation in the experiment was supposed to raise their levels of attention to each others movements, although this is only an assumption and lacks further support.

Another factor is the degree of normalization, or the differences between the individual persons who are shown. Obviously, persons who greatly differ in their body structures are easier to discriminate than persons who appear similar in their outer shapes. In the current study, only the size of the point-light walkers was normalized, to keep the displays as close as possible to the distal stimuli. Study 1 of this work had shown that size is not a relevant factor for person identification from biological motion. Moreover, other studies used either walkers which were similar in size (Cutting & Kozlowski, 1977) or which were normalized with respect to size (Jokisch et al., submitted). Therefore, size-normalization was the method of choice in the current experiment to match the overall size of the participants without changing the natural dynamics inherent to the displays.

In summary, Study 3 did not reveal any evidence for an enhanced sensitivity for self-generated motion patterns over point-light displays depicting familiar or unfamiliar persons. This contradicts the findings of a number of previous studies which showed a better performance for tasks that involved self-generated movements over a variety of domains. However, these studies always required the participants to explicitly indicate whether a movement was self-generated or whether the agent of an event depicted a different person. This was only the case for the last experiment in the current study, and there are various reasons why there was no effect of the type of walker, as was discussed above.
7. General Discussion

Human observers’ ability to identify individual persons from their gait patterns shown as point-light displays was examined in this thesis. Whereas it was known from several studies that observers are able to discriminate individual walking patterns, the factors determining this performance were still unclear. The present work mainly aimed at identifying these factors. Moreover, it contributed to the question whether self-generated gait patterns can be recognized with a greater accuracy than those that were generated by friends or by unknown persons.

The following paragraphs will briefly summarize the findings of the individual studies, discuss their implications for the field of research on biological motion, and point out possible directions of future studies. The final section contains an outlook on potential applications of the findings in this work.

The first two studies investigated the influence of structural and kinematic parameters on the identification of previously unknown walkers by systematically removing information from the displays. Whereas the focus of the first study was mainly on structural properties of the walkers, the second study examined the kinematic parameters in more detail.

Study 1 showed that observers are quickly able to learn to discriminate individual gait patterns from unknown persons shown as point-light displays. These results confirmed the findings of Stevenage et al. (1999) who also demonstrated this ability, but for a mixed-gender group of walker models. Only male walkers were shown in the current study, which presumably increased the difficulty of the task. This study was the first to examine the relevance of individual properties of the models for an identification task. Overall size, body structure, and gait frequency were systematically normalized during the course of the experiment and therefore rendered uninformative for the purpose of differentiation between the walkers. The observers had to rely on a diminishing set of variables to solve the task until only kinematic parameters (i.e., the Fourier components) remained as individual information in the displays.
Whereas it was found that the size of the walkers was not important for their identification, body structure and gait frequency seemed to play a significant role for this task. One could argue that by estimating the effects of these factors from the training sessions, size was not given a “fair chance” to demonstrate its importance, because performance had obviously not saturated when the displays were normalized for size. A different normalization sequence might have put forth a different finding concerning the drop of performance after withdrawing information from the walkers. On the other hand, the results of the training sessions were confirmed by the test sessions that were implemented throughout the experiment. Walkers that were normalized for size could be identified as well as the original displays. It can therefore be concluded that the overall size of a walker does not play an important role in his discrimination from the other displays.

From the point-light displays that were used in Study 1 it cannot be inferred that the size of a person is not used for his or her identification in our everyday environment. On the contrary, individual size (especially body height) can contain an important reference to the identity. On the other hand, this will only work for persons who can be clearly distinguished by their sizes, because observers are unlikely to be able to correctly estimate body sizes which vary only in the scope of a few centimetres. Thus, they often have to rely on different parameters before pictorial cues like face or clothing are discernible. From the results of the first study one can conclude that body structure and gait frequency can be vital and reliable candidates for this purpose.

However, the influence of body shape and gait frequency might have been caused by modifying the apparent dynamics that act on an individual’s walking pattern. These dynamics change in accordance to the structure, because forces act differently on the body. Hence, the kinematic pattern of a walker should change when the body shape is altered, which was not the case with the point-light displays used in this study. Several studies (e.g., Bingham, 1987; Jokisch & Troje, 2003; Runeson & Frykholm, 1981, 1983) have shown that human observers are sensitive for the dynamics inherent in point-light displays. Thus, it is possible that the impaired recognition is based at least in part on the modified dynamics of the displays.
Altering the individual gait frequency might have affected recognition performance in a similar way. Human gait patterns are adjusted with respect to their gait frequencies in order to minimize the required energy (Alexander, 1989; Bertram, 2005; Bertram & Ruina, 2001), which in turn depends on gait speed and step length. An alteration of the gait frequency therefore results in non-optimal energy consumption. The study of Jokisch and Troje (2003) as well as the third experiment of Study 3 indicated that observers are sensitive to naturally adjusted gait frequencies. Thus, this alteration might have had a share in the degraded performance of the observers.

The characteristics of the learning curves of Study 1 support this assumption: after normalization for shape as well as for gait frequency, there is a drop in the levels of recognition in all groups. However, in both cases performance recovers and reaches higher levels after two additional training sessions. Maybe the observers adjusted their criteria to the altered dynamics in the course of these sessions.

Identifying persons shown as isolated point-light displays on a screen is a rather artificial task, because persons always appear in a specific context. To evaluate the importance of the parameters mentioned so far, especially the size, one has to create a more natural setting that contains external references which can be used for comparative purposes. The setting of Jokisch and Troje (2003) might be well suited as a first step in this direction. The authors showed point-light animations of a walking dog in a landscape which contained landmarks to indicate size dimensions. Because the absolute size of the dog was constant throughout the experiment, the apparent size varied with the perceived distance from the observer. Such objects of constant size would be a reference for the size of the walker, and they could be used for estimating the importance of body size in ecological settings.

Nevertheless, it is commonly accepted that the use of point-light animations is a valid method for the extraction of motion components from full-image videos, to examine the role of specific factors in the absence of pictorial cues. So far, many studies have shown observers’ abilities to infer a vast number of valid information categories from point-light displays, for example, gender, emotions, or actions. Moreover, the results from Bülthoff et al. (1998) indicate that these 2D-projections are valid representations of the underlying
3D-models, and the visual system even tolerates depth distortions. Thus, point-light displays of human models can be deemed as viable objects of research, representing the original motions of the distal stimulus. The methods used in this thesis allow for the decomposition of full-body motions of real actors into structural and kinematic components, and the studies presented were the first to apply these methods to the identification of individual walkers. The results thus constitute a valuable contribution to the field of research on biological motion.

Despite the highly degraded displays at the end of the first study, identification performance saturated at a level of 80-90%, depending on the viewpoint. Thus, the kinematic information that remained in the displays was sufficient to allow for the accurate discrimination of individual walking patterns. However, it was not clear whether purely kinematic properties were sufficient to learn to identify the walkers, because structural parameters were still included in the displays during the first training sessions of Study 1. Therefore, only gait patterns which were normalized with respect to size, shape, and gait frequency were used in the second study.

The same seven walkers were used for all observers in Study 1. Thus, it is possible that there were some characteristic patterns among the walkers that were easy to identify, leading to an overestimation of the observers’ abilities to solve the task. An increased database of 40 walkers was available for Study 2, and 20 walkers with a medium distinctiveness were used in the experiments. Therefore, noticeable walking patterns were unlikely to cause enhanced performances. Nevertheless, the results of the second study confirmed the findings of Study 1. Apparently, kinematic properties of walkers are sufficient for their discrimination when they are shown as point-light displays. The learning curves of all three experiments of the first two studies are similar to each other, and performance in both experiments of Study 2 saturated at a level of about 80%, which seems to be a general limit with the degree of normalization used throughout these experiments.

In the first experiment of Study 2 the displays were normalized with respect to their harmonics, that is, only the individual first, second, or third to fifth
harmonics were included in the gait patterns, whereas the remaining harmonics were averaged over the set of walkers. The results revealed significant effects of the first and second harmonics, that is, identification performance dropped significantly when these harmonics were excluded from the displays. On the other hand, performance was still above chance when the walkers were normalized for all but these harmonics. The first two harmonics therefore seem to be sufficient to adequately represent the gait pattern of a walker in a point-light display, although an explicit experiment with walkers that are represented by the first two harmonics has to be conducted to support this assumption. Practical implications of this low-order representation will be discussed in the final part of this chapter.

The second experiment of Study 2 showed that the amplitudes of the dots' trajectories are clearly more important than the phase spectra. The latter only seem to be used from a frontal view, when the individual amplitudes are not as distinguishable from each other as from the other viewpoints. Obviously, in an identification task including walkers of the same sex, the phase spectrum is not of major importance. This might be different in tasks of gender classification. Visual analyses of the displays indicate that the phase spectra of male and female walkers might be different from each other and thus provide a reliable basis for gender discrimination. A prominent feature is the phase relation of the hip- and knee-joints: in female walkers, the vertical motions of these two joints are in counterphase to each other, that is, if one of the hip joints performs a downward motion, the ipsilateral knee is moving upwards. In contrast, the ipsilateral joints of male walkers perform simultaneous vertical motions in the same direction. This might be a reason for the pronounced and sometimes striking hip swing in females.

Kozlowski and Cutting (1977) examined the role of a number of dynamic point-light configurations for the recognition of the sex of the walker. One of these configurations was a combination of the hip-, knee-, and ankle-joints, and the average recognition rate was 55%. This is a rather low value, although it reached significance due to the high number of participants in the study. However, only three walkers were shown to the observers under this condition, and the walkers were shown from a profile view. Gender recognition would probably be more accurate from a frontal view, due to the enhanced access to
the phase relations of the joints. A normalization procedure as used throughout the current work could be applied to examine the role of the phase spectrum for gender recognition, in order to ascertain whether this phase relation is an indicator of a particularly male or female gait pattern.

The findings of Studies 1 and 2 showed that the best recognition performance was found when training and test viewpoints were identical. This is in accordance with results from Verfaillie (1993) who found that only an identical viewpoint between priming and primed stimulus had a beneficial effect on the reaction time in a recognition experiment. However, the present results also suggest that there is a substantial generalization from one viewpoint to the others, although performance clearly declined after a rotation of the viewing angle between training and test sessions. Even under maximum rotation the identification rate reached about 40%, which is noticeably above chance level of 14%. Obviously, observers do not only learn to discriminate absolute motion trajectories that are projected onto the screen, but they also create an abstract knowledge of how these point-light displays appear from a different perspective. Tests of viewpoint generalization were always based on walkers that were normalized for size, shape, and gait frequency. Thus, the remaining kinematic properties are sufficient to create these representations.

Performance declined with an increasing degree of rotation between training and test viewpoints, with a major impairment already observable after a rotation to or from the half-profile view. It is unclear, however, whether performance would deteriorate even more under rotation angles that exceeded the ones used in the present studies. Most important, it would be instructive to examine identification rates under symmetric viewpoints, for example, under half-profile views that are rotated to the left and to the right, respectively. In a study about face perception, Troje and Bülthoff (1998) found that symmetric views resulted in an enhanced recognition of the faces, compared to otherwise different viewpoints. However, this viewpoint advantage was obviously caused by an image-based matching of the stimuli, and not by a matching of the bilaterally symmetric spatial features. This was concluded from the finding that mirrored (and thus unrealistic) faces could be recognized with an even better accuracy than the realistic stimuli, seen from symmetric viewpoints.
Walking patterns, on the other hand, are less symmetric than static faces, because people often show a pronounced lateral sway to either side, or a different amount of arm swing. One would therefore assume a lower sensitivity for the recognition of walkers from symmetric viewpoints than this is the case for face matching. However, whereas the effect of recognizing faces from different viewpoints has been examined in a large number of studies, to my knowledge this approach has not been applied to biological motion yet.

It is not clear what kind of representation the ability to generalize from one viewpoint to another is based upon. The fact that there is a transfer between viewing angles only means that observers have a concept of how the motions of the point-lights depicted from one viewing angle might look from different viewpoints. They do not necessarily have an idea about the underlying persons who generated these motions, and how these persons would move in a different context. It would be instructive to apply different types of test stimuli to examine the concept that is used in the recognition process. For example, with the use of modern graphics software, it is possible to create fully animated 3D-characters whose motions are driven by the same markers that are used to compute the 15 virtual markers of the displays shown in the present experiments. A good recognition performance would indicate a more general representation of the model, beyond the basic point-light trajectories. Another possibility would be to use point-light depictions of the same persons performing different kinds of motions. This would be an even stronger test of the concept that observers have when they recognize individual motion patterns. In this case, recognition performance above chance level would suggest that the observers have an idea about the “acting style” of a person, and of the underlying dynamic principles that act on an individual’s motion pattern.

Still another way to test for the generalization would be to use slightly different versions of the models’ gait patterns. From the methods used to capture the data and to create the trajectories of the virtual markers, it is highly improbable that artefacts remained in the displays, for example, a “twitch” of one of the point-lights that indicated the presence of a specific person. Averaging over a number of gait cycles removes accidental deviations from the normal gait patterns from the final displays. Nevertheless, to test for this
possibility, one could create two versions of each walker’s point-light representation, which might slightly differ from each other, due to the nature of the data. Transfer from one dataset to the other should be about as good as that found in the current studies.

Another interesting question is the one about the underlying “walker space”, that is, the space in which all walkers are represented and in which they are compared to each other. After decomposing the gait patterns, each walker is mathematically represented as a single point in a multidimensional space. Presumably, gait patterns that exhibit large distances from each other in this space also vastly differ in their visual appearances, and they can thus be easily discriminated. On the other hand, walking patterns that look similar are supposedly based on representations that are in close vicinity in this space. By relating an individual motion pattern to a reference, for example, the average of a group of walkers, it is possible to exaggerate the differences between these two values. This is often referred to as “caricaturing”.

In the domain of face recognition, these caricatures have often been found to facilitate recognition by enhancing the distinctiveness of the faces, but nevertheless retaining their identities (e.g., Benson & Perrett, 1994; Lee, Byatt, & Rhodes, 2000; Rhodes, Brennan, & Carey, 1987). A caricature represents an extrapolation of the vector joining the average face to a particular individual, and manipulating the position of the face on this vector changes the face’s individuality (Leopold, O’Toole, Vetter, & Blanz, 2001). It was argued that veridical faces may be stored as exaggerated versions, and caricatures therefore match these representations closer than they match the original faces. Another explanation is that exaggerations provide a better access to the memorized versions (Rhodes et al., 1987). Leopold et al. (2001) used an adaptation paradigm to examine the visual representations underlying face recognition. They found that adaptation to individual faces biased the subsequent identification of other faces. More specifically, recognition was facilitated for faces on the trajectory passing through the adaptive and average faces, so called “anti-faces”, with respect to the adaptive face.

Hill and Pollick (2000) showed that temporal exaggerations of point-light displays of arm movements lead to a better recognition of these movements
than their veridical versions. In principle, this should also be the case for whole-body movements, and the walker representation used in this study provides a valuable instrument for the examination of this question. By using an exaggeration- or adaptation-paradigm, it would be possible to examine whether there is a specific point of reference that the individual motion patterns are compared against. For example, in the second study this could be the average of the seven walkers that were used for each respective subject. Another possible reference could be a prototypical gait pattern that is created by averaging a large group of walkers, including either only males or walkers of both sexes. Comparing recognition performances for point-light stimuli that are created by exaggerating veridical stimuli against varying reference points could give a hint on the relative importance of these references and indicate whether motion patterns are structured around a specific point in the perceptual space.

The first two studies examined whether it is possible to learn to discriminate previously unknown locomotion patterns from point-light displays, and which factors contribute to this ability. In contrast, Study 3 focused on the role of the familiarity with a gait pattern for a number of perceptual tasks. These tasks included the recognition of a human walker in a scrambled mask, the discrimination between normal and reversed displays, and the adjustment of the original gait frequency of walkers. In case of a significant role of motor experience with a gait pattern, as postulated by the common coding approach (Prinz, 1997), self-generated motion patterns should lead to higher performances in these tasks than those of friends or strangers. Temporal and spatial characteristics of action effects ought to reflect the characteristics of an individual's action plannings, which vary between persons. On the other hand, if visual experience had an effect on the performance, this should have become evident in an enhanced performance when the observers were presented point-light displays of people known to them. Whereas the findings on the apparent motion direction indicated a general effect of visual familiarity with gait patterns, none of the experiments provided evidence for a significant role of motor or visual experience with specific gait patterns for the tasks that were used.

This result might have been caused by the tasks themselves. Previous studies that demonstrated a facilitation of the recognition of self-generated
actions (e.g., Knoblich & Flach, 2001; Knoblich & Prinz, 2001; Loula et al., 2004) required the observers to discriminate or to name the agent of an action. In contrast, the current experiments applied indirect tasks, without the agents of the actions being known to the observers. Apparently, these tasks were not sufficient to activate representations of the subjects’ action effects.

The first experiment of Study 3 showed that the point-light walkers that are created by applying motion capture and a subsequent decomposition of the data are highly salient, even under visual noise of 120 additional dots. Many studies that can be found throughout the literature still use adapted versions of Cutting’s original algorithm (1977), although this obviously lacks ecological validity, because it is not based on a natural gait pattern, but was created by estimating structural and kinematic parameters. It seems reasonable that the methods used in this work are an excellent way to provide realistic stimuli for studies on the perception of biological motion. Moreover, the representation of the walkers enables experimenters to change the relevant parameters with little effort and to create stimuli that fit the individual’s needs. However, a direct comparison of the saliency of this representation and the algorithm provided by Cutting (1977) should be conducted to further support these assumptions.

Finally, let’s take an outlook on possible applications. It was shown that relatively complex gait patterns can be physically decomposed into a low-order Fourier series with hardly losing any relevant information, that is, a big portion of the overall variance is covered by the first two harmonics. On the other hand, this representation leads to a percept of individuality. Observers are able to discriminate individuals based on these depictions. It seems reasonable that future projects can rely on these low-order representations of point-light walkers for research on the perception of individuality from biological motion. A limited set of harmonics means that fewer variables have to be processed in the computation of the displays. A walker model that is represented by five harmonics contains a set of 497 variables which have to be processed for each frame that is presented on the screen. Discarding the third to fifth harmonics, which do not seem to play significant roles in the perception of an individual’s gait pattern, would remove 270 variables from the necessary computations, leading to a vastly enhanced performance. This in turn raises the probability
that each frame can be displayed on the screen, allowing for a maximum of control over the stimuli that are presented to the observers. On the other hand, it would release resources of the computer to display additional stimuli which are of interest for the respective research project, for example, flankers or masking dots.

However, creating realistic individual motion patterns is not only a fascinating topic to be covered by research. There is a growing digital entertainment industry, with animated characters becoming increasingly realistic. Some of these characters are modelled on real persons, for example, the cowboy Woody in *Toy Story*, whose motions were inspired by the movements of Tom Hanks. Knowing which features of a motion pattern lead to the perception of an individual would enable digital artists to concentrate on these aspects. Another branch of interest is the video gaming industry. Many games that include human motion as an important factor (e.g., sports games) rely on common locomotion patterns which are applied to every character in the same way. A slight alteration of the kinematic parameters, however, could lead to an enhanced individuality of these characters and to an increased perception of uniqueness. With the linear methods applied in the creation of the gait patterns it is even possible to examine the perception of other properties, for example, traits or moods. Troje (2002b) showed that these attributes can be reliably capture and visualized in a vivid way. This possibility should be used to establish an even higher degree of personality with the characters. Of course, a gait pattern is only one possible movement to be modelled, but since it is the single most encountered human motion pattern, it would be a good point to start at.
References


References


**Lebenslauf**

**Persönliche Daten**

Name: Cord Westhoff  
Geburtstag: 7. März 1975  
Geburtsort: Bad Oeynhausen

**Ausbildung**

1981 - 1985: Grundschule Vlotho  
1985 - 1994: Weser-Gymnasium Vlotho; Abitur  
1994 - 1995: Zivildienst am Wittekindshof Bad Oeynhausen (Einrichtung für Geistigbehinderte)  
1995 - 2002: Studium der Psychologie an der Universität Bielefeld  
1997: Vordiplom Psychologie  
2001: Diplom-Arbeit: „Hemisphärenaktivierung bei Schlaganfallpatienten unter visuo-perzeptiver Stimulation“ (Blutflussmessung durch Doppler-Sonographie)  
Betreuer: Dr. Michael Finger  
Feb. 2002: Diplom Psychologie

**Wissenschaftliche Erfahrung**

Apr. - Dez. 2002: Wissenschaftlicher Mitarbeiter am Institut für Medizinische Psychologie (Leiter: Prof. Dr. B. Sabel), Otto-von-Guericke-Universität Magdeburg; Schwerpunkt auf Beeinträchtigungen nach Schädigungen des visuellen Systems

seit Jan. 2003: Wissenschaftlicher Mitarbeiter im BioMotionLab (Leiter: Prof. Dr. Nikolaus Troje), Institut für Kognitive Neurowissenschaften, Ruhr-Universität Bochum; Erforschung der Wahrnehmung menschlicher Bewegung durch den Einsatz von Motion Capture; Titel der Dissertation: „Person Identification from Biological Motion“
Danksagung

Folgenden Personen möchte ich herzlichst für die Unterstützung bei der Fertigstellung der Dissertation danken:

- Herrn Prof. Nikolaus Troje, der die Entstehung dieser Arbeit betreut hat und durch seine Ideen und Anregungen wesentlich an ihrer Vollendung beteiligt war.

- Herrn Prof. Onur Güntürkün, der sich als Gutachter für diese Arbeit zur Verfügung gestellt hat.

- Tobias Otto, mit dem ich ungezählte unterhaltsame Stunden auf zwei verschiedenen Kontinenten verbracht habe, und der durch freundschaftliche und technische Unterstützung unschätzbare Beiträge geleistet hat (das war schon okay!).

- Susanne Bergert, die wertvolle Vorschläge zur Auswertung und inhaltlichen Gestaltung beigetragen hat.

- Allen Korrekturlesern, die durch konstruktive Beiträge an der Verbesserung der Inhalte beteiligt waren.

- Nicht zuletzt meinen Eltern, die mich während meiner kompletten Ausbildung bedingungslos unterstützt haben, und die mir meinen Weg bis hierhin in dieser Form erst ermöglichten.